To require the Federal Trade Commission to identify content-agnostic platform interventions to reduce the harm of algorithmic amplification and social media addiction on covered platforms, and for other purposes.

IN THE SENATE OF THE UNITED STATES

Ms. KLOBUCHAR (for herself and Ms. LUMMIS) introduced the following bill; which was read twice and referred to the Committee on

A BILL

To require the Federal Trade Commission to identify content-agnostic platform interventions to reduce the harm of algorithmic amplification and social media addiction on covered platforms, and for other purposes.

1 Be it enacted by the Senate and House of Representa-
2 tives of the United States of America in Congress assembled,

3 SECTION 1. SHORT TITLE.

4 This Act may be cited as the “Nudging Users to
5 Drive Good Experiences on Social Media Act” or the “So-
6 cial Media NUDGE Act”.

7 SEC. 2. FINDINGS.

8 Congress finds the following:
(1) Social media platforms can have significant impacts on their users, both positive and negative. However, social media usage can be associated with detrimental outcomes, including on a user’s mental and physical health. Design decisions made by social media platforms, such as decisions affecting the content a user might see on a social media platform, may drive or exacerbate these negative or detrimental outcomes.

(2) Viral harmful content often spreads on social media platforms. Social media platforms do not consistently enforce their terms of service and content policies, leading to supposedly prohibited content often being shown to users and amplified by such platforms.

(3) Social media platforms often rely heavily on automated measures for content detection and moderation. These social media platforms may rely on such automated measures due to the large quantity of user-generated content on their platforms. However, evidence suggests that even state-of-the-art automated content moderation systems currently do not fully address the harmful content on social media platforms.
(4) Significant research has found that content-agnostic interventions, if made by social media platforms, may help significantly mitigate these issues. These interventions could be readily implemented by social media platforms to provide safer user experiences. Such interventions include the following:

(A) Nudges to users and increased platform viewing options, such as screen time alerts and grayscale phone settings, which may reduce addictive platform usage patterns and improve user experiences online.

(B) Labels and alerts that require a user to read or review user-generated content before sharing such content.

(C) Prompts to users, which may help users identify manipulative and microtargeted advertisements.

(D) Other research-supported content-agnostic interventions.

(5) Evidence suggests that increased adoption of content-agnostic interventions would lead to improved outcomes of social media usage. However, social media platforms may be hesitant to independently implement content-agnostic interventions that
will reduce negative outcomes associated with social
media use.

SEC. 3. STUDY ON CONTENT-AGNOSTIC INTERVENTIONS.

(a) Study to Identify Content-agnostic Interven-
tions.—The Director of the National Science Foun-
dation (in this section referred to as the “Director”) shall
enter into an agreement with the National Academies of
Sciences, Engineering, and Medicine (in this section re-
ferred to as the “Academies”) to conduct ongoing studies
to identify content-agnostic interventions that covered
platforms could implement to reduce the harms of algo-
rithmic amplification and social media addiction on cov-
ered platforms. The initial study shall—

(1) identify ways to define and measure the
negative mental or physical health impacts related to
social media, including harms related to algorithmic
amplification and social media addiction, through a
review of—

(A) a wide variety of studies, literature, re-
ports, and other relevant materials created by
academic institutions, civil society groups, and
other appropriate sources; and

(B) relevant internal research conducted
by a covered platform or third party research in
the possession of a covered platform that is vol-
untarily submitted to the Academies by the covered platform (through a process, established by the Academies, with appropriate privacy safeguards);

(2) identify research-based content-agnostic interventions, such as reasonable limits on account creation and content sharing, to combat problematic smartphone use and other negative mental or physical health impacts related to social media, including through a review of the materials described in subparagraphs (A) and (B) of paragraph (1);

(3) provide recommendations on how covered platforms may be separated into groups of similar platforms for the purpose of implementing content-agnostic interventions, taking into consideration factors including any similarity among the covered platforms with respect to—

(A) the number of monthly active users of the covered platform and the growth rate of such number;

(B) how user-generated content is created, shared, amplified, and interacted with on the covered platform;

(C) how the covered platform generates revenue; and
(D) other relevant factors for providing recommendations on how covered platforms may be separated into groups of similar platforms;

(4) for each group of covered platforms recommended under paragraph (3), provide recommendations on which of the content-agnostic interventions identified in paragraph (2) are generally applicable to the covered platforms in such group;

(5) for each group of covered platforms recommended under paragraph (3), provide recommendations on how the covered platforms in such group could generally implement each of the content-agnostic interventions identified for such group under paragraph (4) in a way that does not alter the core functionality of the covered platforms, considering—

(A) whether the content-agnostic intervention should be offered as an optional setting or feature that users of a covered platform could manually turn on or off with appropriate default settings to reduce the harms of algorithmic amplification and social media addiction
on the covered platform without altering the
core functionality of the covered platform; and

(B) other means by which the content-agnostic intervention may be implemented and
any associated impact on the experiences of
users of the covered platform and the core
functionality of the covered platform;

(6) for each group of covered platforms recom-
ommended under paragraph (3), define metrics gen-
erally applicable to the covered platforms in such
group to measure and publicly report in a privacy-
-preserving manner the impact of any content-agnostic
intervention adopted by the covered platform;
and

(7) identify data and research questions nec-
essary to further understand the negative mental or
physical health impacts related to social media, in-
cluding harms related to algorithmic amplification
and social media addiction.

(b) Requirement to Submit Additional Re-
search.—If a covered platform voluntarily submits inter-

nal research to the Academies under subsection (a)(1)(B),
the covered platform shall, upon the request of the Acad-
emies and not later than 60 days after receiving such a
request, submit to the Academies any other research in
the platform’s possession that is closely related to such voluntarily submitted research.

(c) **Reports.**—

(1) **Initial study report.**—Not later than 1 year after the date of enactment of this Act, the Academies shall submit to the Director, Congress, and the Commission a report containing the results of the initial study conducted under subsection (a), including recommendations for how the Commission should establish rules for covered platforms related to content-agnostic interventions as described in paragraphs (1) through (5) of subsection (a).

(2) **Updates.**—Not later than 2 years after the effective date of the regulations promulgated under section 4, and every 2 years thereafter during the 10-year period beginning on such date, the Academies shall submit to the Director, Congress, and the Commission a report containing the results of the ongoing studies conducted under subsection (a). Each such report shall—

(A) include analysis and updates to earlier studies conducted, and recommendations made, under such subsection;

(B) be based on—
(i) new academic research, reports, and other relevant materials related to the subject of previous studies, including additional research identifying new content-agnostic interventions;

(ii) new academic research, reports, and other relevant materials about harms occurring on covered platforms that are not being addressed by the content-agnostic interventions being implemented by covered platforms as a result of the regulations promulgated under section 4;

(C) include information about the implementation of the content-agnostic interventions by covered platforms and the impact of the implementation of the content-agnostic interventions; and

(D) include an analysis of any entities that have newly met the criteria to be considered a covered platform under this Act since the last report submitted under this subsection.

SEC. 4. IMPLEMENTATION OF CONTENT-AGNOSTIC INTERVENTIONS.

(a) Determination of Applicable Content-agnostic Interventions.—
10

(1) IN GENERAL.—Not later than 60 days after the receipt of the initial study report under section 3(c)(1), the Commission shall initiate a rulemaking proceeding for the purpose of promulgating regulations in accordance with section 553 of title 5, United States Code—

(A) to determine how covered platforms should be grouped together;

(B) to determine which content-agnostic interventions identified in such report shall be applicable to each group of covered platforms identified in the report; and

(C) to require each covered platform to implement and measure the impact of such content-agnostic interventions in accordance with subsection (b).

(2) CONSIDERATIONS.—In the rulemaking proceeding described in paragraph (1), the Commission—

(A) shall consider the report under section 3(c)(1) and its recommendations; and

(B) shall not promulgate regulations requiring any covered platform to implement a content-agnostic intervention that is not discussed in such report.
11

(3) Notification to covered platforms.—

The Commission shall, not later than 30 days after the promulgation of the regulations under this subsection, provide notice to each covered platform of the content-agnostic interventions that are applicable to the platform pursuant to the regulations promulgated under this subsection.

(b) Implementation of Content-agnostic Interventions.—

(1) In general.—

(A) Implementation plan.—

(i) In general.—Not later than 60 days after the date on which a covered platform receives the notice from the Commission required under subsection (a)(3), the covered platform shall submit to the Commission a plan to implement each content-agnostic intervention applicable to the covered platform (as determined by the Commission) in an appropriately prompt manner. If the covered platform reasonably believes that any aspect of an applicable intervention is not technically feasible for the covered platform to implement, would substantially change the core functionality
of the covered platform, or would pose a
material privacy or security risk to con-
sumer data stored, held, used, processed,
or otherwise possessed by such covered
platform, the covered platform shall in-
clude in its plan evidence supporting these
beliefs in accordance with paragraph (2).

(ii) COMMISSION DETERMINATION.—
Not later than 30 days after receiving a
covered platform’s plan under clause (i),
the Commission shall determine whether
such plan includes details related to the
appropriately prompt implementation of
each content-agnostic intervention applica-
able to the covered platform, except for any
aspect of an intervention for which the
Commission determines the covered plat-
form is exempt under paragraph (2).

(iii) APPEAL OR REVISED PLAN.—

(I) IN GENERAL.—Subject to
subclause (II), if the Commission de-
determines under clause (ii) that a cov-
ered platform’s plan does not satisfy
the requirements of this subsection,
not later than 90 days after the
issuance of such determination, the covered platform shall—

(aa) appeal the determination by the Commission to the United States Court of Appeals for the Federal Circuit; or

(bb) submit to the Commission a revised plan for a Commission determination pursuant to clause (ii).

(II) LIMITATION.—If a covered platform submits 3 revised plans to the Commission for a determination pursuant to clause (ii) and the Commission determines that none of the revised plans satisfy the requirements of this subsection, the Commission may find that the platform is not acting in good faith in developing an implementation plan and may require the platform to implement, pursuant to a plan developed for the platform by the Commission, each content-agnostic intervention applicable to the platform (as determined by the Com-
mission) in an appropriately prompt manner.

(B) Statement of compliance.—Not less frequently than annually, each covered platform shall make publicly available on their website and submit to the Commission, in a machine-readable format and in a privacy-preserving manner, the details of—

(i) the covered platform’s compliance with the required implementation of content-agnostic interventions; and

(ii) the impact (using the metrics defined by the Director of the National Science Foundation and the National Academies of Sciences, Engineering, and Medicine pursuant to section 3(a)(6)) of such content-agnostic interventions on reducing the harms of algorithmic amplification and social media addiction on covered platforms.

(2) Feasibility, functionality, privacy, and security exemptions.—

(A) Statement of inapplicability.—

Not later than 60 days after the date on which a covered platform receives the notice from the
Commission required under subsection (a)(3), a covered platform seeking an exemption from any aspect of such rule may submit to the Commission—

(i) a statement identifying any specific aspect of a content-agnostic intervention applicable to such covered platform (as determined by the Commission under subsection (a)) that the covered platform reasonably believes—

(I) is not technically feasible for the covered platform to implement;

(II) will substantially change the core functionality of the covered platform; or

(III) will create a material and imminent privacy or security risk to the consumer data stored, held, used, processed, or otherwise possessed by such covered platform; and

(ii) specific evidence supporting such belief, including any relevant information regarding the core functionality of the covered platform.
(B) **Determination by the Commission.**—Not later than 30 days after receiving a covered platform’s statement under subparagraph (A), the Commission shall determine whether the covered platform shall be exempt from any aspect of a content-agnostic intervention discussed in the covered platform’s statement.

(C) **Appeal or Revised Plan.**—Not later than 90 days after a determination issued under subparagraph (B), a covered platform may—

(i) appeal the determination by the Commission to the United States Court of Appeals for the Federal Circuit; or

(ii) submit to the Commission a revised plan, including details related to the prompt implementation of any content-agnostic intervention for which the covered platform requested an exemption that the Commission subsequently denied, for a Commission determination pursuant to paragraph (1)(A)(ii).
SEC. 5. TRANSPARENCY REPORT.

Not later than 180 days after the date of enactment of this Act, and semiannually thereafter, each covered platform shall publish a publicly-available, machine-readable report about the content moderation efforts of the covered platform with respect to each language spoken by not less than 100,000 monthly active users of the covered platform in the United States. Such report shall include the following:

(1) CONTENT MODERATORS.—The total number of individuals employed or contracted by the covered platform during the reporting period to engage in content moderation for each language, broken down by the number of individuals retained as full-time employees, part time employees, and contractors of the covered platform and reported in a privacy-preserving manner.

(2) RANDOM SAMPLE OF VIEWED CONTENT.—Information related to a random sample of publicly visible content accounting for 1,000 views each month. Each month, covered platforms shall calculate the total number of views for each piece of publicly visible content posted during the month and sample randomly from the content in a manner such that the probability of a piece of content being sampled is proportionate to the total number of views of
that piece of content during the month. Covered platforms shall report the following information about each piece of sampled content (with appropriate redactions to exclude the disclosure of illegal content):

(A) The text, images, audio, video, or other creative data associated with each such piece of content.

(B) The details of the account or accounts that originally posted the content; and

(C) The total number of views of each such piece of content during the month

(3) HIGH REACH CONTENT.—Content moderation metrics broken down by language to assess the prevalence of harmful content on the covered platform, including, for each language, the 1,000 most viewed pieces of publicly visible content each month, including the following (with appropriate redactions to exclude the disclosure of illegal content):

(A) The text, images, audio, video, or other creative data associated with each such piece of content.

(B) The details of—

(i) the account that originally posted the content; and
(ii) any account whose sharing or re-posting of the content accounted for more than 5 percent of the views of the content.

(4) REMOVED AND MODERATED CONTENT.—

(A) IN GENERAL.—Aggregate metrics for user-generated content that is affected by any automated or manual moderation system or decision, including, as calculated on a monthly basis and reported in a privacy-preserving manner, the number of pieces of user-generated content and the number of views of such content that were—

(i) reported to the covered platform by a user of the covered platform;

(ii) flagged by the covered platform by an automated content detection system;

(iii) removed from the covered platform and not restored;

(iv) removed from the covered platform and later restored; or

(v) labeled, edited, or otherwise moderated by the covered platform following a user report or flagging by an automated content detection system.
(B) Requirements for metrics.—The metrics reported under subparagraph (A) shall be broken down by—

(i) the language of the user-generated content;

(ii) the topic of the user-generated content, such as bullying, hate speech, drugs and firearms, violence and incitement, or any other category determined by the covered platform to categorize such content; and

(iii) if the covered platform has a process for publicly verifying that an account on the platform belongs to a prominent user or public figure, whether the creator of the content is a politician or journalist with a verified account.


(a) Unfair or Deceptive Acts or Practices.—A violation of section 3(b), 4, or 5 or a regulation promulgated under section 4 shall be treated as a violation of a rule defining an unfair or deceptive act or practice prescribed under section 18(a)(1)(B) of the Federal Trade Commission Act (15 U.S.C. 57a(a)(1)(B)).

(b) Powers of the Commission.—
(1) IN GENERAL.—The Commission shall enforce this Act in the same manner, by the same means, and with the same jurisdiction, powers, and duties as though all applicable terms and provisions of the Federal Trade Commission Act (15 U.S.C. 41 et seq.) were incorporated into and made a part of this Act.

(2) PRIVILEGES AND IMMUNITIES.—Any person who violates section 4 or 5 or a regulation promulgated under section 4 shall be entitled to the privileges and immunities provided in the Federal Trade Commission Act (15 U.S.C. 41 et seq.).

(3) ENFORCEMENT GUIDELINES AND UPDATES.—Not later than 1 year after the date of enactment of this Act, the Commission shall issue guidelines that outline any policies and practices of the Commission related to the enforcement of this Act in order to promote transparency and deter violations of this Act. The Commission shall update the guidelines as needed to reflect current policies, practices, and changes in technology, but not less frequently than once every 4 years.

(4) AUTHORITY PRESERVED.—Nothing in this Act shall be construed to limit the authority of the Commission under any other provision of law.
SEC. 7. DEFINITIONS.

In this Act:

(1) Algorithmic Amplification.—The term “algorithmic amplification” means the promotion, demotion, recommendation, prioritization, or de-prioritization of user-generated content on a covered platform to other users of the covered platform through a means other than presentation of content in a reverse-chronological or chronological order.

(2) Commission.—The term “Commission” means the Federal Trade Commission.

(3) Content Moderation.—The term “content moderation” means the intentional removal, labeling, or altering of user-generated content on a covered platform by the covered platform or an automated or human system controlled by the covered platform, including decreasing the algorithmic ranking of user-generated content, removing user-generated content from algorithmic recommendations, or any other action taken in accordance with the covered platform’s terms of service, community guidelines, or similar materials governing the content allowed on the covered platform.

(4) Content-Agnostic Intervention.—The term “content-agnostic intervention” means an action that can be taken by a covered platform to alter
a user’s experience on the covered platform or the user interface of the covered platform that does not—

(A) rely on the substance of user-generated content on the covered platform; or

(B) alter the core functionality of the covered platform.

(5) COVERED PLATFORM.—The term “covered platform” means any public-facing website, desktop application, or mobile application that—

(A) is operated for commercial purposes;

(B) provides a forum for user-generated content;

(C) is constructed such that the core functionality of the website or application is to facilitate interaction between users and user-generated content; and

(D) has more than 20,000,000 monthly active users in the United States for a majority of the months in the previous 12-month period.

(6) PRIVACY-PRESERVING MANNER.—The term “privacy-preserving manner” means, with respect to a report made by a covered platform, that the information contained in the report is presented in a manner in which it is not reasonably capable of
being used, either on its own or in combination with other readily accessible information, to uniquely identify an individual.

(7) **User.**—The term “user” means a person that uses a covered platform, regardless of whether that person has an account or is otherwise registered with the platform.

(8) **User-generated content.**—The term “user-generated content” means any content, including text, images, audio, video, or other creative data that is substantially created, developed, or published on a covered platform by any user of such covered platform.