



California Terms of Service Report

December 2023 (reporting period July 1 – September 30, 2023)

LinkedIn is a real-identity online service for professionals to connect and interact with other professionals, learn, hire, and find jobs. LinkedIn's vision is to create economic opportunity for every member of the global workforce. Its mission is to connect the world's professionals to make them more productive and successful. As part of that mission, LinkedIn is committed to keeping its platform and services safe, trusted, and professional, and to providing transparency to its members, the public, and to regulators.

LinkedIn provides the information below in response to California Business & Professions Code Section 22677.

1. The current version of the terms of service.

The current versions of LinkedIn's terms of service, as defined under Section 22676(b), are available at the following links:

- **LinkedIn User Agreement:** <https://www.linkedin.com/legal/user-agreement>.
- **LinkedIn Professional Community Policies:** <https://www.linkedin.com/legal/professional-community-policies>.
- **How we enforce our Professional Community Policies:** <https://www.linkedin.com/help/linkedin/answer/a1342754>.
- **How to report inappropriate content, messages, or safety concerns:** <https://www.linkedin.com/help/linkedin/answer/a1378278>.

2. If the company has filed its first report, a complete and detailed description of any changes to the terms of service since the previous report.

This report is LinkedIn's first report; as a result, there are no relevant changes to disclose.

3. A statement of whether the current version of the terms of service define each of the following categories of content, and, if so, the definition of those categories: (A) Hate speech or racism; (B) Extremism or radicalism; (C) Disinformation or misinformation; (D) Harassment; (E) Foreign political interference.

All content on LinkedIn must comply with LinkedIn's [Professional Community Policies](#), which set out in detail the content LinkedIn permits and does not permit to keep its platform safe, trusted, and professional. LinkedIn's Professional Community Policies address each of the categories enumerated in Section 22677(a)(3), as outlined below.

Hate speech and racism are addressed within the [Hateful and derogatory content policy](#), which reads as follows:

Do not be hateful. We don't allow content that attacks, denigrates, intimidates, dehumanizes, incites or threatens hatred, violence, prejudicial or discriminatory action against individuals or groups because of their actual or perceived race, ethnicity, national origin, caste, gender, gender identity, sexual orientation, religious affiliation, age, or disability status. Hate groups are not permitted on LinkedIn. Do not use racial, religious, or other slurs that incite or promote hatred, or any other content intended to create division. Do not post or share content that denies a well-documented historical event such as the Holocaust or slavery in the United States.

LinkedIn provides additional detail regarding the policy and examples of content that violates the policy in the following article, which is linked in the Professional Community Policies:
<https://www.linkedin.com/help/linkedin/answer/a1339812>.

Extremism and radicalism are addressed within the [Dangerous organizations and individuals policy](#), which reads as follows:

Do not share content promoting dangerous organizations or individuals. We don't allow any terrorist organizations or violent extremist groups on our platform. And we don't allow any individuals who affiliate with such organizations or groups to have a LinkedIn profile. Content that depicts terrorist activity, that is intended to recruit for terrorist organizations, or that threatens, promotes, or supports terrorism in any manner is not tolerated.

LinkedIn provides additional detail regarding the policy and examples of content that violates the policy in the following article, which is linked in the Professional Community Policies:
<https://www.linkedin.com/help/linkedin/answer/a1342780>.

Misinformation, disinformation, and foreign political interference are addressed within the [False and misleading content policy](#), which reads as follows:

Do not share false or misleading content: Do not share content that is false, misleading, or intended to deceive. Do not share content to interfere with or improperly influence an election or other civic process. Do not share content that directly contradicts guidance from leading global health organizations and public health authorities; including false information about the safety or efficacy of vaccines or medical treatments. Do not share content or endorse someone or something in exchange for personal benefit (including personal or family relationships, monetary payment, free products or services, or other value), unless you have included a clear and conspicuous notice of the personal benefit you receive and have otherwise complied with our Advertising Policies.

LinkedIn provides additional detail regarding the policy and examples of content that violates the policy in the following article, which is linked in the Professional Community Policies:
<https://www.linkedin.com/help/linkedin/answer/a1340752>.

Harassment is addressed in the [Harassment](#) and [Sexual harassment policies](#) (together, "Harassment"), which read as follows:

Do not post harassing content. We don't allow bullying or harassment. This includes targeted personal attacks, intimidation, shaming, disparagement, and abusive language directed at other members. Do not reveal others' personal or sensitive information (for example, doxing), or

incite others to do the same. You may not engage in trolling or other repetitive negative content that disrupts other members or conversations on the platform.

Do not engage in sexual innuendos or unwanted advances. We don't allow unwanted expressions of attraction, desire, requests for romantic relationship, marriage proposals, sexual advances or innuendo, or lewd remarks. LinkedIn is a professional networking platform, not a dating site. Do not use LinkedIn to pursue romantic connections, ask for romantic dates, or provide sexual commentary on someone's appearance or perceived attractiveness. Do not send unwanted advances in messages, posts, or comments or send sexually explicit images to anyone on the platform.

LinkedIn provides additional detail regarding the policies and examples of content that violates the policies in the following articles, which are linked in the Professional Community Policies: <https://www.linkedin.com/help/linkedin/answer/a1342756>; <https://www.linkedin.com/help/linkedin/answer/a1337770>.

4. Description of content moderation practices.

As noted above, all content on LinkedIn must comply with LinkedIn's [Professional Community Policies](#), which set out in detail the content LinkedIn permits and does not permit to keep its platform safe, trusted, and professional.

LinkedIn applies a three-layer, multidimensional approach to moderate content on LinkedIn:

- The first layer of protection is automated and proactive prevention. When a member attempts to create a piece of content on LinkedIn, various calls (or signals) are sent to LinkedIn's machine learning services. These services aim to automatically filter out certain policy-violating content at the time of creation.
- The second layer of protection is a combination of automated and human-led detection. LinkedIn's second layer of moderation detects content that's likely to be violative but for which LinkedIn is not sufficiently confident to warrant automatic removal, and sends it for human review.
- The third layer of protection is human-led detection. If users locate content that they believe violates LinkedIn's policies, they are able to report it for review using LinkedIn's in-product reporting functionality.

A. Any existing policies intended to address the categories of content described in Business & Professions Code Section 22677(a)(3).

Please see the response in Section 3 above.

B. How automated moderation systems enforce terms of service of the platform and when these systems involve human review.

LinkedIn is a global platform with over 1 billion members worldwide and utilizes both human content moderation and machine-learned automated systems to moderate content. LinkedIn's automated systems help enforce its terms of service in three relevant ways. First, LinkedIn's automated systems in some cases determine content violates LinkedIn's policies, and action that content. Second, LinkedIn's automated systems may queue content that potentially violates LinkedIn's policies for human review

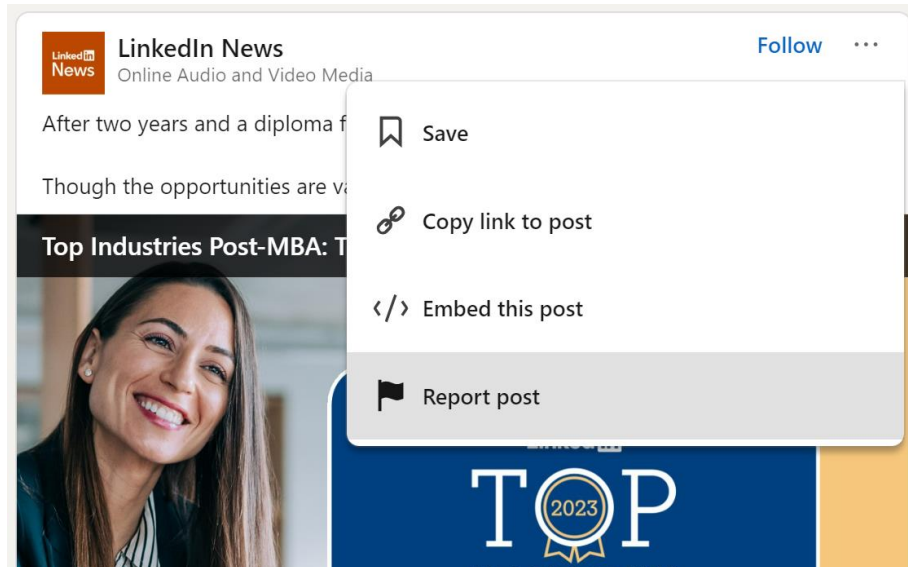
and decision. Third, LinkedIn’s automated systems in some cases evaluate and resolve user reports of policy-violating content, and in other cases queue user reports for human review and decision.

LinkedIn employs the following safeguards, among others, to its automated systems for content moderation: LinkedIn’s monitors the aggregate performance and accuracy of its moderation systems, and sets minimum thresholds for performance; LinkedIn sets thresholds for individual decisions made by the systems, such that the system will not act on a piece of content or user report and will instead send it for human review; LinkedIn controls the types of reports the system acts on (e.g., the system does not automatically resolve reports of terrorist content); LinkedIn generally allows authors to appeal a decision if they believe the decision is incorrect; and LinkedIn periodically retrains its system to account for, for example, changes in human-reviewer decisions, content trends, and user-report trends over time.

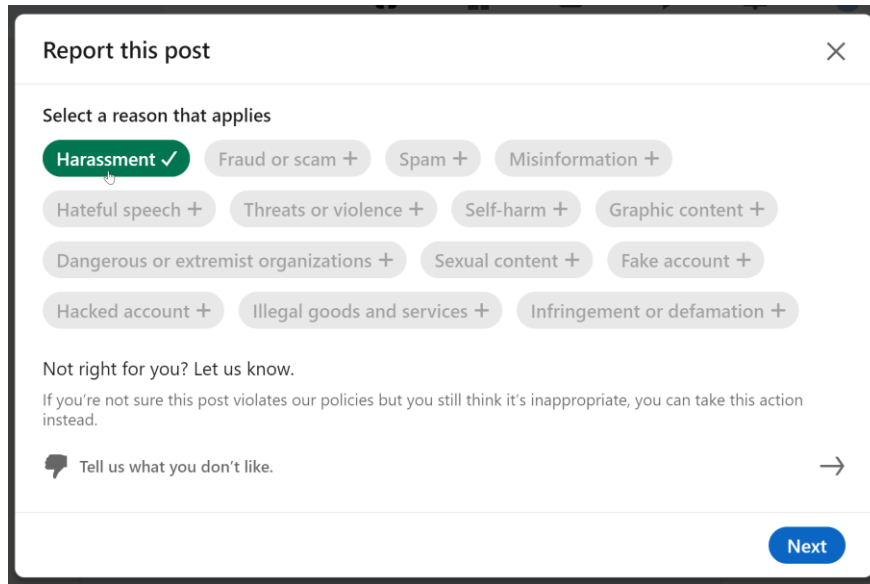
C. How the company responds to user reports of violations of the terms of service.

To report content, members click the three-dot icon available in-product on the content and follow the in-product prompts. For example, to report a post on LinkedIn, members use the following process:

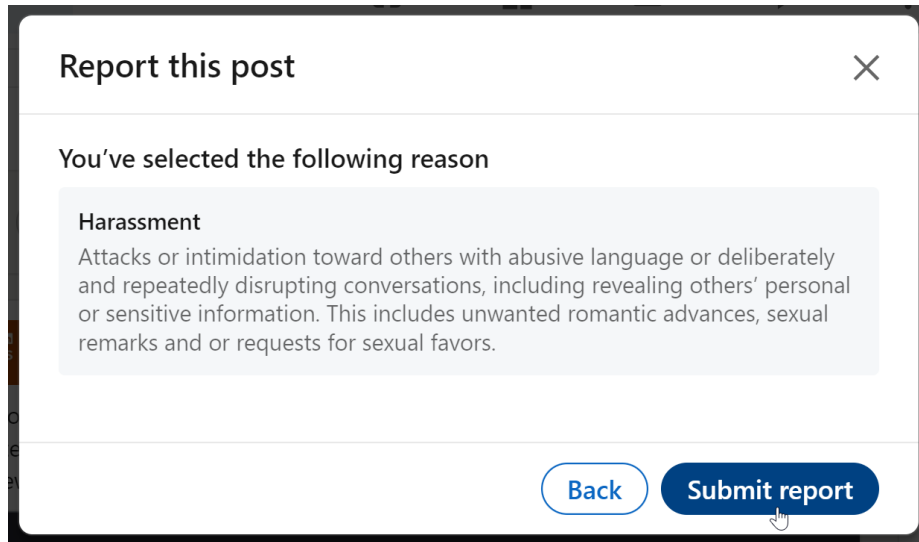
1. Select the three-dot icon in the upper-right corner of the post, and select ‘Report post’:



2. Select the reporting reason that applies to the post. For example, “Harassment”:



3. Review the selected reporting reason, and submit the report:



When users report content, those reports are sent for review to determine whether the content in question violates LinkedIn’s policies. LinkedIn typically notifies the reporting user of its decision. If a violation is found, LinkedIn removes or actions the content, consistent with its policies.

D. How the company would remove individual pieces of content, users, or groups that violate the terms of service, or take broader action against individual users or against groups of users that violate the terms of service.

Violating LinkedIn’s policies can result in action against a piece of content (including, for example, a post, a page, a group, etc.) as well as an account. When we remove content, we also generally send the member a notice that the content violates LinkedIn’s policies, how the content violates our policies, and

the action we've taken. If a member believes their content has been removed in error, they have the opportunity to submit an appeal.

Repeated violations can result in account restriction. We provide the opportunity to appeal an account restriction, and we may reinstate a restricted account if the member agrees to comply with the Professional Community Policies. Continued violations result in permanent restriction from the LinkedIn platform.

For certain egregious violations of our policies (e.g., child sexual abuse material, terrorism, egregious sexual harassment), we may permanently restrict a member account after a single violation.

In addition, LinkedIn maintains an internal Trust & Safety team staffed with investigators who proactively investigate coordinated abuse (for example, groups of accounts engaged in scam activity). When the Trust & Safety team identifies violations of LinkedIn's policies, they action the content and accounts consistent with the process above.

E. The languages in which the platform does not make terms of service available, but does offer product features, including but not limited to, menus and prompts.

LinkedIn's platform is available in twenty-six supported languages, outlined in the Help Center article here: [LinkedIn Supported Languages](#). The terms of service are available in each of these languages.

5. Relevant content moderation metrics.

In response to Section 22677(a)(5), the sections below provide metrics regarding the content moderation actions LinkedIn took worldwide during the reporting period (July 1 – September 20, 2023) for the four policies outlined in Section 3 above: Hateful and derogatory content; Dangerous organizations or individuals; False and misleading content; and Harassment (together, "relevant policies"). The metrics reported are worldwide metrics across LinkedIn's platform and not geographically limited. Note that LinkedIn also actions content under its other policies – for example, copyrighted content, adult content, scam content, etc. – outside the scope of Section 22677(a)(3) and not included in the metrics below.¹

During the period, LinkedIn applied two relevant actions to content because it violated LinkedIn's policies:

1. Action 1: LinkedIn removed content that violated its policies; and
2. Action 2: LinkedIn limited the visibility of content that violated its policies.

A. Number of flagged items of content.

The table below reports the number of flagged items of content during the reporting period, disaggregated by relevant policy. The counts include both content flagged by users as well as content flagged and actioned by LinkedIn's reviewers and automated system.

¹ The metrics LinkedIn provides in this report are best estimates provided the data available in LinkedIn's systems and methods used in the ordinary course of business. In some cases, metrics can be impacted by, for example, account deletion, content deletion, as well as downtime or errors in LinkedIn's systems that may impact data recording. Certain data may also vary or change over time. Metrics in the report are based on data as of the close of the reporting period. These metrics are prepared for the purposes of Cal. Bus. & Prof. Code Section 22677 and may not be suitable for other purposes.

LinkedIn Policy	Number of flagged items of content ²³
Hateful and derogatory	285,945
Dangerous organizations or individuals	12,564
False and misleading	212,235
Harassment	69,378

B. Number of actioned items of content.

The table below reports the number of items of content LinkedIn actioned under the relevant policies during the reporting period, including the action LinkedIn applied, disaggregated by policy.

LinkedIn Policy ⁴	Number of actioned items of content	Number of items of content where LinkedIn applied Action 1	Number of items of content where LinkedIn applied Action 2
Hateful and derogatory	196,104	196,104	0
Dangerous organizations or individuals	24	24	0
False and misleading	76,644	57,352	19,292
Harassment	14,749	14,749	0

C. Number of views, viewers, and reshares of the actioned items of content.

The table below reports the approximate number of logged-in views, viewers, and reshares of the actioned content in Table 5(B) above during the reporting period, disaggregated by policy.

LinkedIn Policy	Number of views	Number of viewers	Number of reshares
Hateful and derogatory	17,365,675	6,856,069	13,474
Dangerous organizations or individuals	2,621	2,530	8
False and misleading	49,282,300	16,750,314	71,150
Harassment	2,562,033	1,560,983	3,623

² A single piece of content can be reported by multiple users for differing report reasons. For example, a single piece of content may be flagged by one user as ‘Hateful’ content and another as ‘Harassment.’ In such a case, the table above separately includes the content in both the ‘Hateful and derogatory’ and ‘Harassment’ counts. As a result, the counts may overstate the number of unique flagged pieces of content on LinkedIn.

³ Except where otherwise noted, ‘content’ addressed in this report includes user-generated content that appears in LinkedIn’s Feed – for example, posts, articles, pages, groups, comments, newsletters, etc. – as well as job posts that appear on LinkedIn’s jobs board. The metrics do not include, e.g., messages, accounts/profiles, or ads. LinkedIn looks forward to expanding its metrics coverage in future reports.

⁴ The LinkedIn policy categories are broader than the specific harms listed in Cal. Bus. & Prof. Code Section 22677(a)(3). For example – as described in Section 3 above – LinkedIn’s ‘Hateful and derogatory content’ policy, and the counts in the table above, include not only instances of “Hate speech or racism” but also, e.g., a job post with age discrimination in violation of LinkedIn’s policy. Accordingly, the counts above may overstate Section 22677(a)(3) content on LinkedIn.

D. Number appeals of actioned items of content.

The table below reports the number of author appeals LinkedIn received during the reporting period for the actioned content, from Table 5(B) above, as well as the number of appeals LinkedIn granted.

LinkedIn Policy	Number of appeals	Number of appeals granted
Hateful and derogatory	2,987	166
Dangerous organizations or individuals	0	0
False and misleading	214	11
Harassment	315	2

E. Number of accounts restricted.

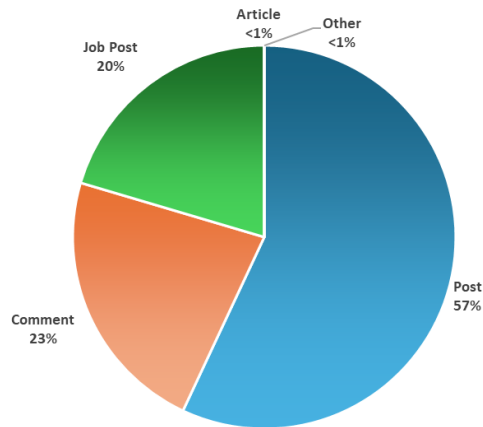
The table below reports the number of member accounts LinkedIn permanently restricted during the reporting period because the account repeatedly posted content that violated one of the relevant policies.

As described above, in some cases LinkedIn may permanently restrict an account after a single egregious content policy violation (e.g., in the case of egregious sexual harassment). Such restrictions are included in the count below. The count below does not include temporary restrictions, nor does it include restrictions unrelated to the relevant policies (e.g., permanent restrictions for copyright infringement, permanent restrictions for posting content in violation of LinkedIn’s scam or child exploitation policies, permanent restrictions for data scraping, etc.).

Permanent account restrictions	263
---------------------------------------	-----

F. Metric disaggregations.

Actioned items of content, by content type

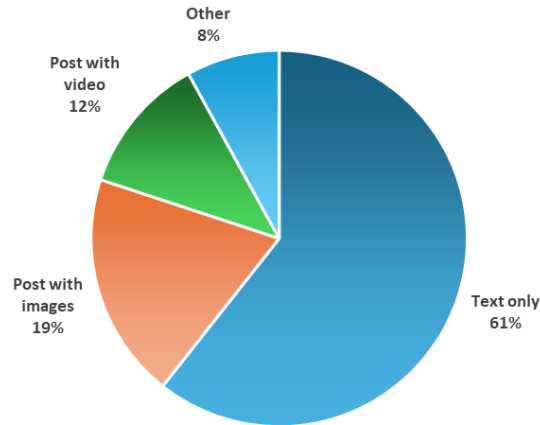


The table below reports actioned items of content, from Table 5(B) above, disaggregated by content type (i.e., post, comment, article, job post, or other).

Content type	Count
Post	163,832
Comment	65,105

Article	27
Job post	58,541
Other (e.g. Pages, Groups, Events)	16

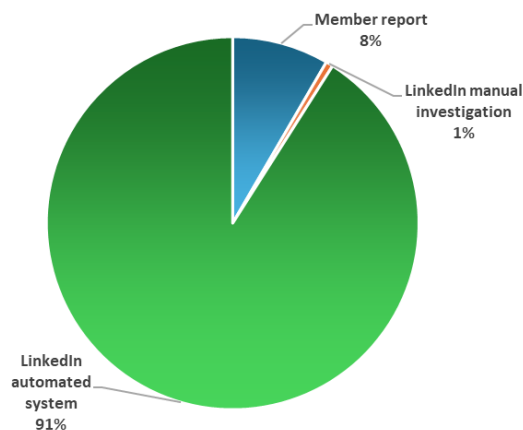
Actioned items of content, by media type



The table below reports actioned items of content, from Table 5(B) above, disaggregated by media type (i.e., text only, post with images, post with video, or other).

Media type	Count
Text only	174,371
Post with images	55,861
Post with video	34,284
Other / miscellaneous	23,055

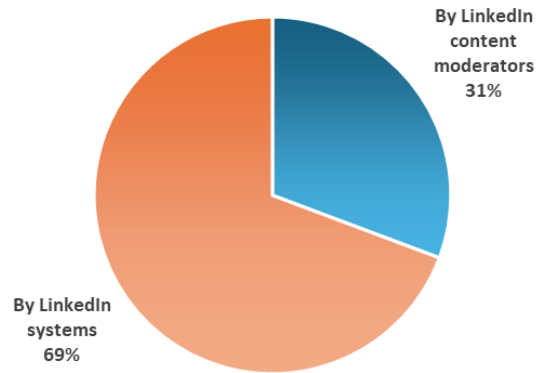
Actioned items of content, by how the content was flagged



The table below reports actioned items of content, from Table 5(B) above, disaggregated by how the content was flagged (i.e., member report, LinkedIn manual investigation, or LinkedIn automated system).

How the content was flagged	Count
Member report	24,124
LinkedIn manual investigation	1,791
LinkedIn automated system	261,606

Actioned items of content, by how the content was actioned



The table below reports actioned items of content, from Table 5(b) above, disaggregated by how the content was actioned (i.e., by LinkedIn content moderators or by LinkedIn systems).

How the content was actioned	Count
By LinkedIn content moderators	88,275
By LinkedIn systems	199,246