

Snap Inc.

紐約服務條款報告

2025 年 7 月 1 日至 9 月 30 日



提交日期: 2026 年 1 月 1 日

紐約服務條款報告(2025年7月1日至9月30日)

Snap Inc.

Snap Inc. 依據《紐約州一般商業法》第 42 篇第 1100–1104 條之規定，向紐約州檢察長提交本服務條款報告。本紐約服務條款報告涵蓋期間為 2025 年 7 月 1 日至 2025 年 9 月 30 日(2025 年第 3 季)。

我們的條款(紐約州一般商業法第 1101、1102 條)

我們努力地為 Snapchat 上的創意與自我表達打造安全又有趣的環境。所有 Snapchat 用戶都必須遵守我們的[服務條款](#)，包括我們的[社群規範](#) (在本報告中統稱為「條款」)。有關我們如何審核內容與執行政策的其他資訊，請參閱我們的[社群規範](#)說明系列，其中包含我們的[審核、執行與申訴](#)政策的描述，以及有關每個類別內容與行為的其他資訊。

除條款外，我們還在[安全中心](#)提供安全相關資訊與資源，包括有關如何在服務中[檢舉涉嫌違反](#)條款或其他安全問題的指導。

上述所有文件皆以英文附於本報告中，並已在我們的網站上提供紐約地區英語能力有限者最常使用的 12 種非英語語言版本(涵蓋我們提供 Snapchat 服務的語言)。

內容審核政策與實務(紐約州一般商業法第 1102 條)

我們的條款禁止紐約州一般商業法第 1102(c) 條所列之內容類型，說明如下：

內容類別	社群規範 禁止的相應內容類別	相關定義與政策，請參閱我們的 透明度報告詞彙表 與 社群規範 解釋系列。
仇恨言論或種族主義	仇恨言論(屬於仇恨內容、恐怖主義與暴力極端主義)	仇恨言論或內容基於種族、膚色、種姓、民族、國籍、宗教、性向、性別認同、殘疾或退伍軍人身分、移民身分、社會經濟地位、年齡、體重或懷孕狀況，貶低、誹謗或鼓勵歧視或暴力。這些規則禁止種族、族群、厭女症或仇視同性戀的誹謗內容。並且禁止嘲笑或呼籲歧視受保護團體的迷因，以及故意使用舊名錯稱或性別錯稱的迷因。仇恨言論也包括認可人類悲劇(如種族滅絕、種族隔離或奴隸制度)的加害者或貶低受害者。其他禁止的仇恨內容包括使用仇恨標誌，指任何為在表達對他人仇恨或歧視的意象。如需更多資訊，請詳閱我們 關於仇恨內容、恐怖主義和暴力極端主義的說明 。
極端主義或激進化	恐怖主義與暴力極端主義(屬於仇恨內容、恐怖主義與暴力極端主義)。	宣揚個人或團體，為實現意識型態目標而實施恐怖主義或其他暴力犯罪行為的內容。上述規則也禁止任何宣揚或支持外國恐怖組織或暴力極端主義仇恨團體(由可信的第三方專家判斷)，及招募此類組織或暴力極端主義活動的內容。如需更多資訊，請詳閱我們 關於仇恨內容、恐怖主義和暴力極端主義的說明 。
虛假或錯誤資訊	錯誤資訊(屬於有害虛假或欺騙行為)。	包括造成傷害或帶有惡意的錯誤或誤導性內容，例如否認悲劇事件存在、未經證實的醫學聲明、破壞公民程序的公正性，或基於錯誤或誤導性目的經過操縱的內容(包括透過生成式 AI 或欺騙性的編輯)。如需更多資訊，請詳閱我們關於 有害或欺騙行為的說明 。
騷擾	(1)騷擾與霸凌，以及(2)性騷擾(屬於性內容範疇)(以下合稱為「騷擾」，如以下圖表所示)。	係指任何可能使一般人產生情緒困擾的不受歡迎行為，例如言語辱罵、威脅，或任何意圖使人羞辱、難堪或貶低的行為。如需更多資訊，請詳閱我們關於 騷擾與霸凌的說明 。 此外，我們在「性內容」相關的《社群規範》中，亦全面禁止各種形式的性騷擾。這可能包含作出不受歡迎的挑逗行為、分享影像與未經允許的內容，或向其他使用者傳送猥褻的要求或邀請。如需更多資訊，請詳閱我們關於 色情內容的說明 。

外國政治干預	不實資訊(屬於有害的不實或欺騙性行為, 包含但不限於外國或政治來源)。	如需瞭解我們對錯誤資訊的定義, 請參閱上文。此類別亦包含冒充行為, 即帳戶假冒與其他個人或品牌有關聯。如需更多資訊, 請詳閱我們關於 有害或欺騙行為的說明 。
--------	-------------------------------------	---

我們的[審核、執行與申訴說明](#)與[嚴重有害說明](#)提供以下主題的詳細資訊:

- 我們如何透過自動化工具與人工審核內容、
- 我們如何回應用戶檢舉涉嫌違反社群規範的行為, 以及
- 我們如何對違反《社群規範》的個別內容與用戶進行處置。

違反我們條款的相關資訊(2025年7月1日至9月30日)(紐約州一般商業法第1102條)

以下提供在2025年7月1日至9月30日期間, 透過應用程式內回報或由系統自動偵測之違反《社群規範》的詳細資訊, 並符合紐約州一般商業法第1102條之規定。當中提供的資料為全球性。下列數據不包含透過Snapchat應用程式以外(例如支援網站與電子郵件)提交的檢舉, 此類檢舉量占總檢舉量不到1%。

除另有規定外, 本節中使用的術語均根據我們的[透明度詞彙表](#)定義。

違反事宜類別	行為警告標記	已標記的內容或帳戶總數 ⁽¹⁾	由人工審核人員執行的內容 ⁽²⁾	由自動化工具強制執行的內容	由人工審核人員執行的不重複帳戶 ⁽³⁾	由自動化工具強制執行的不重複帳戶	針對由人工審核員執行的帳戶鎖定所提出的申訴	針對由自動化工具強制執行的帳戶鎖定提出的申訴	申訴後恢復的帳戶 ⁽⁴⁾ (最初由人工審核人員鎖定)	申訴後恢復的帳戶 (最初由自動化工具鎖定)	人工審核人員強制執行的內容的暴力瀏覽率 (VVR) ⁽⁵⁾	VVR 內容, 由自動化工具強制執行	由人工審核者強制執行的內容下不重複的暴力觀看率 ⁽⁶⁾	由自動化工具強制執行的內容下不重複的違反觀看率
仇恨言論	人類報告	482,240	144,389	20,679	123,835	18,343	94	0	5	0	0.000205%	0.000004%	0.41%	0.01%
	主動偵測	3,073	2,441	32	2,029	29	7	1,164	1	15	0.000002%	0.00000%	0.00%	0.00%
恐怖主義與暴力極端主義	人類報告	199,245	559	54	468	54	199	0	1	0	0.000037%	0.00000%	0.07%	0.00%
	主動偵測	12,240	5,901	0	3,754	0	11	38	1	6	0.00000%	0.00000%	0.00%	0.00%
假訊息	人類報告	386,210	145	593	113	593	1	0	0	0	0.000003%	0.00000%	0.01%	0.00%
	主動偵測	262	10	0	10	0	1	0	0	0	0.00000%	0.00000%	0.00%	0.00%
騷擾	人類報告	3,184,734	756,736	604,798	624,142	491,972	700	0	17	0	0.001135%	0.000168%	1.93%	0.35%
	主動偵測	12,293	8,972	1,583	7,670	1,428	17	13,442	1	97	0.00007%	0.00000%	0.02%	0.00%

- (1) 因可能違反社群規範的內容或帳戶總數, 包括向我們檢舉的內容或帳戶總數, 以及透過我們的主動偵測流程通報的內容或帳戶總數。為了將這些資料分解為違規內容的類別, 我們使用了採取執法行動的最終執法原因。如果內容或帳戶受到標記但未採取執法行動, 我們會將指標歸因於受標記內容或帳戶的疑似違反類別。
- (2) Snapchat 上已處置內容 (例如 Snap 和故事) 的數量。「處置」是指對內容或帳戶採取的行動 (例如刪除、警告和鎖定)。
- (3) Snapchat 上已處置的不重複帳戶數量。例如, 如果單一帳戶由於各種原因遭到多次處置 (例如, 使用者因發表錯誤資訊遭到警告, 然後因為騷擾其他使用者使得帳戶被鎖定), 這個統計資料只會有一個帳戶受處置。如上所述, 「處置」是指對內容或帳戶採取的行動 (例如刪除、警告、鎖定)。
- (4) 我們只會復原我們的版主認定被錯誤鎖定的帳戶。
- (5) 違規內容收視率是包含違規內容之故事與 Snap 的觀看次數, 佔 Snapchat 上所有故事和 Snapchat 觀看次數的比率。(Snap 是使用 Snapchat 相機拍攝的照片或影片。如需更多資訊, 請參閱[這裡](#)。)舉例而言, 如果 VVR 為 0.03%, 表示 Snapchat 上每 10,000 次 Snap 和故事觀看次數, 有 3 次違反我們政策的內容。這項指標讓我們瞭解, Snapchat 上的 Snap 與故事觀看次數, 有多少百分比是觀看違反社群規範的內容 (遭到檢舉或被我們強制執行的內容)。
- (6) 不重複違規內容收視率是指在整個報告期間 (即 2025 年第三季) 內, 不重複活躍用戶中, 看到違規故事與/或 Snap 的不重複觀眾所佔的比例。舉例而言, 如果我們的不重複違規內容收視率為 0.03%, 代表在相關期間 Snapchat 上每 10,000 名活躍用戶, 就有 3 名觀眾觀看了違反我們政策的故事和/或 Snap。此指標讓我們瞭解 Snapchat 上的用戶中有多少比例會違反我們的社群規範 (已檢舉或主動執行) 的故事與/或 Snap。

其他資訊

如需進一步瞭解 Snap 的安全、隱私與透明度, 請造訪我們的[隱私、安全與政策中心](#), 以及我們的[透明度報告頁面](#)。