Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

《傳播與社會學刊》，（總）第71期（2025）：183–209
https://doi.org/10.30180/CS.202501_(71).0008

專輯論文

中國語境下平台內容審查的權力及其界限——以隱形禁止為例

張璐[a]

[a]中國政法大學光明新聞傳播學院，北京市，中國大陸

摘要

平台基於一定公共屬性、技術壟斷和先天優勢，在與平台用户之間支配與被支配關係中形成了一種典型的「私權力」，塑造了有組織的內容審查「私人秩序」。平台內容審查的舉措及其影響是本文研究的中心問題。本研究綜合應用了文本分析、深度訪談和問卷調查等方法，對中國平台內容審查，尤其是「隱形禁止」（shadow banning）措施的現實作出調研。研究發現中國既有的法律法規為平台採取「隱形禁止」措施提供了寬鬆環境和有利條件；平台內容審查的私權力已呈現侵害使用者權利的現實風險。本文從規範用户協議、算法、技術應用、程序公開透明等方面提出完善中國平台內容審查權力的建議。

關鍵詞：平台治理、內容審查、平台責任、隱形禁止、私權力

張璐，中國政法大學光明新聞傳播學院助理教授。研究興趣：傳播法、平台責任、個人資訊保護、人工智能治理。電郵：zhanglu@cupl.edu.cn
論文投稿日期：2024年2月26日。論文接受日期：2024年7月25日。

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

Communication and Society, *71* (2025), 183–209

Special Issue Article

The Power and Limits of Platform Content Moderation in the Chinese Context: Taking Shadow Banning as an Example

Lu ZHANG[a]

[a] Guangming School of Journalism and Communication, China University of Political Science and Law, Beijing, Mainland China

Abstract

Based on certain public attributes, technological monopolies, and inherent advantages, platforms have established a form of private power in the dominant–subordinate relationship with their users, thus shaping an organized private order of content censorship. The measures and impacts of platform content censorship are the central issues examined in this article. This study employed text analysis, in-depth interviews, and surveys to investigate content censorship on Chinese platforms, focusing especially on shadow banning measures. It found that existing laws and regulations in China create a lenient environment and favorable conditions for platforms to adopt shadow banning measures. The private power of platform content censorship has already shown a real risk of infringing on users' rights. This article offers suggestions for improving the power of content censorship on Chinese platforms in terms of regulating user agreements, applying algorithmic technology, and ensuring procedural transparency.

Keywords: platform governance, content censorship, platform responsibility, shadow banning, private authority

Lu ZHANG (Assistant Professor). Guangming School of Journalism and Communication, China University of Political Science and Law. Research interests: communication law, platform responsibility, personal information protection, regulation of artificial intelligence. Email: zhanglu@cupl.edu.cn

Article History: Received on 26 February 2024. Accepted on 25 July 2024.

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

引言

數字時代，平台重塑經濟生產的過程和組織樣態，改變信息傳播方式和影響，成為政治權力和市場權利之間的第三力量(張晨穎，2021)。其中，平台對非法內容的規制採用信息篩選和內容呈現等多重方式，這被認為是體現國家公權力、平台私權力、用戶私權力複雜博弈最為充分的場景(孔祥穩，2020)。一方面，依託平台展開規制成為世界範圍內的普遍選擇，各國立法都對平台在非法信息的傳播中施加了一定的注意義務；另一方面，平台基於一定公共權力、技術壟斷和先天優勢，在規制非法內容的時候容易濫用權力，對用戶權利造成不當干預。作為追求商業利益最大化的市場主體，平台勢力日益強大的同時也更容易引發權力濫用和失範風險，對平台自我治理的治理是基於平台自治權產生的一個新命題(張新平，2023)。當前，中國以《網絡安全法》、《電子商務法》和《互聯網信息服務管理辦法》等法律法規為基礎，輔以大量的規章和規範性文件，構建起一個以平台為中心的規制體系。但是面對平台治理侵權和失範風險的增加，現有的平台治理法律法規對此鮮有規範，這一問題近年來已引起了多位學者的關注(如趙鵬，2016；姚志偉，2019；孔祥穩，2020；劉權，2020；張新平，2023)。在此背景下，本文擬以「隱形禁止」(shadow banning)這一新型平台內容審查為切入，探討平台內容審查的規範現狀和實踐情況，重點分析平台侵蝕使用者權利的可能失範風險，並提出約束平台內容審查權力的建議，旨在推動網絡平台在法治軌道上健康安全繁榮發展。

「隱形禁止」：平台內容審查的新形式

隱形禁止的本體論

對在線內容進行審查的傳統典型方式是刪除內容和封禁賬號，這被Goldman (2021) 稱之為「保留或刪除」的「二元範式」。當前，刪除內容和封禁賬號仍是平台內容審查的重要手段。除此之外，隱形禁止日益成為平台內容管理的其他重要措施，是平台為限制內容可見性而採取的秘密行動(Gillespie, 2022)。「隱形禁止」是一種口語化的表述，由一個叫作「Something Awful」的網站和論壇的版主於2001年創造，最初

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

指在網絡論壇上的一種具有欺騙性質的賬號封禁方式，即被禁止的用戶不知道其發佈的內容已經不再被其他用戶可見（Radsch, 2021）。2018年，時任美國總統的特朗普發表推文，指責Twitter利用「隱形禁止」來審查保守派觀點，這一事件讓隱形禁止獲得了更多的關注（Savolainen, 2022）。

「隱形禁止」可以作為動詞（意為以秘密方式禁止用戶），也可以用作指代這種禁令的名詞，還可以是形容詞（「被禁止的」）。2022年，歐盟通過《數字服務法》（*The Digital Services Act*），首次在法律層面直接使用「隱形禁止」術語。雖然法案正文並沒有出現「隱形禁止」，但是法案基於第55條條款（Recital 55）明確了「隱形禁止」的含義，將其視為「限制可見性」（restriction of visibility）的一種類型，涵蓋各種阻止用戶內容被廣泛地流通、降低被意外看到（being accidentally viewed）可能性的現象。[1]

一般而言，我們可以在兩個維度上使用隱形禁止。一種是狹義的隱形禁止，僅限於秘密地「刪除內容或封禁賬號」的行為，強調與傳統公開封禁賬號和內容不同，只要是用戶不知情的（未被通知的）封禁行為，均屬於「隱形禁止」。另一種是廣義的隱形禁止，除了秘密地「刪除內容或封禁賬號」之外，還包括降低特定內容可見性的行為，比如通過算法等技術實行內容排序、推薦系統中降低權重、影響特定內容可見性的所有行為。廣義上的隱形禁止行為範圍非常廣，反映了平台在內容治理中使用算法和人工智能技術的趨勢。本文將在廣義上使用該術語。

隱形禁止的複雜性

隱形禁止的術語爭議背後，折射了其本身的複雜性。作為隱藏或限制用戶內容而不通知用戶其發生的行為，隱形禁止背後存在多元利益的糾葛。其一，隨著平台私權力的擴大，平台不僅有不得發佈、傳播非法內容的法定義務，還附有防止用戶非法內容擴散的法定義務。依託平台展開在線內容治理成為世界範圍內的普遍選擇，附有監管職責的行政部門推動了平台責任認定的嚴格化和平台義務的擴張化。因為基於數據權力、技術壟斷和先天優勢，不論是從平台自身利益而

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

言，還是從成本效率的角度來看，平台都處於內容管理的最佳位置。由此，平台為了合規義務，自然會在實踐中積極採取包括隱形禁止在內的多元內容審查措施。

其二，平台將封禁賬號等內容管理措施視為其聲譽風險和用戶問責的源頭，而可見性限制的措施將有利於減輕這種風險，維護用戶對平台的信任關係。內容管理與言論自由之間的矛盾與衝突是每個平台都棘手的問題，為了營造正面的企業形象和良好的線上言論環境，各個平台都不太可能直接承認自己在內容審查方面的措施。平台通常面臨著兩難選擇：如果平台輕易採取刪除、屏蔽等措施，會導致平台的用戶黏性降低，直接影響到平台的商業利益；但是如果平台不採取措施，則可能需要就擴大的損害與用戶承擔連帶責任。在這種情況下，可見性降低的內容管理方式具有很大的適用空間，尤其是對待那些合法但是令人不適的內容，或者遊走於法律邊緣的在線內容。降低內容可見性較封禁、屏蔽等手段成為平台管理用戶內容中更為「左右逢源」的手段。

其三，平台之所以選擇不通知目標用戶，與披露成本高、排序的複雜性等也有緊密關係。從實際角度來看，平台對內容和用戶進行治理的系統是複雜的、技術性的和多層次的。平台很難通過簡潔明瞭的語言向用戶進行有意義的披露和說明。雖然有時用戶可以對內容管理提出申訴，然後將問題內容傳送給人工審核員進行評估和覆核。然而，更普遍的情況是，算法要麼直接自動刪除內容，要麼將內容標記給人工審核員。人類的考量和算法的參與以越來越難以辨別的方式交織在一起，即使平台想要更加開放地解釋其內容呈現的規則，但是現實的技術和認識上的障礙會限制其可以披露的程度。

此外，內容管理和排序在一定程度上是平台為用戶提供的一種優化反饋循環的先進工具。通過內容管理和排序，用戶關注他們的內容表現如何，並相應地進行修改，由此可以指導普通用戶朝著與數字營銷相關的行為和實踐方向發展。當然，也存在這樣一種設想，平台之所以不披露信息呈現的規則，是擔心被用戶利用規則。因此最好讓用戶對規則是甚麼、如何執行、是否違反了規則，甚至是否因此受到懲罰等，都保持一種不確定性的狀態。如果用戶收到了封禁賬號或者內

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

容管理的通知，他們將會選擇創建新賬號，或者通過試錯等反向「操作」企圖完美逃避檢測（Cotter, 2019）。當然，法律對於在線非法內容的標準和舉措其實是相對模糊的，在程序正當上也缺少規定，這些給予平台選擇秘密地降低可見性等措施提供了廣闊的空間。

中國平台內容審查的規範分析

中國對在線內容審查具有較多規範和豐富實踐。關於平台內容審查的規定最早見於2000年實施的《互聯網信息服務管理辦法》。該辦法已於2011年修改，但是最重要的第15條，即對內容審查的「九不准」標準並沒有特別修訂。「九不准」對中國在線非法內容審查實踐確立了基準，多年來產生了重要影響。近年來，隨著算法等信息技術的不斷發展，中國先後頒佈了多部法律法規對在線內容審查做出積極回應，這些規範中既有專門針對內容審查的特別規範（比如2020年《網絡信息內容生態治理規定》等），也有在對其他技術做出約束的同時提及在線內容審查條款（比如2019年頒佈的《互聯網信息服務算法推薦管理規定》第6條強調不得利用算法推薦服務傳播法律、行政法規禁止的信息等）。既有規範全面覆蓋了新聞信息服務、即時通信、信息搜索和直播等各領域的平台媒介，並提出「平台主體責任」的法律義務，要求平台對所托管的內容負責。

平台非法內容審查的相關規定

各國對平台內容審查的規範分析都聚焦於兩個關鍵問題：（一）非法信息的判定標準為何？（二）平台對於在線非法內容有何義務？

合理、準確判定非法信息的範圍，是實現在線內容審查的前提和基礎。關於在線內容非法性的判斷標準，最詳細的要屬上文提到的「九不准」。2000年頒佈的《互聯網信息服務管理辦法》第15條規定了九種非法內容的信息，包括反對憲法所確立的基本原則：危害國家安全、洩露國家秘密、顛覆國家政權、破壞國家統一、損害國家榮譽和利益等。[2]《互聯網信息服務管理辦法》是國務院頒佈的行政法規，從效力層

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

級上來看，行政法規雖不及法律，「九不准」對在線內容審查卻產生了重要影響，此後中國的多部法律法規在對內容判斷的標準上基本是沿用了「九不准」的標準。但是即使是以細緻著稱的「九不准」，在具體使用時仍面臨著因規範過於抽象和模糊而帶來的適用難題，比如「九不准」包括「損害國家榮譽和利益」的信息，但是何為「損害國家榮譽和利益」則缺乏具體明確的判定規則，導致實踐中難以直接適用。此外，在對非法內容認定的條文中，法律法規往往採用「列舉＋兜底」的規範格式，呈現以「其他非法信息」為兜底條款的粗放型設計特點，這樣進一步擴大了對非法內容判斷的自由裁量空間。

此外，關於平台對非法內容的義務要求上，早期中國法律對平台內容審查的義務主要集中於事後處置措施，包括停止傳輸、保存記錄、報告有關機關三項義務。近年來相關法律法規新增了包括消除信息、制止刪除、防止信息擴散等處置義務，《互聯網信息服務算法推薦管理規定》還增加了不得屏蔽、過度推薦、操縱檢索結果排序等干預信息呈現的義務。從內容審核流程的角度來看，停止傳輸、保存記錄、報告有關機關等義務側重發現非法內容之後採取的措施，而屏蔽、推薦等干預信息呈現義務則覆蓋信息傳播的全流程，從這個角度表明了平台責任嚴格化的趨勢。

與隱形禁止相關的規定

各國的法律雖然對內容審查都有或多或少的規定，但是在隱形禁止上卻很少有直接的條文。歐盟《數字服務法》因在條款第55條使用了「shadow ban」的術語而被認為是世界上第一部直接針對該舉措的法律（Leerssen, 2023）。但是值得注意的是，不論術語和概念為何，以隱形禁止為代表，平台對內容管理卻是不爭的事實，由此引發關於平台內容管理的失範問題、約束平台權力是全世界共同的關注。本文認為隱形禁止的合法性問題有兩個關鍵要點：第一，隱形禁止是一種調節內容排序的手段，因而平台是否具有管理內容呈現順序的權力是一個前提問題。第二，隱形禁止最大的特徵在於隱蔽性，即用戶並不知道自己發佈內容的可見度降低。因而，如果法律規定平台採取處置措施必須「告

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

知用戶」的話，那麼，隱形禁止的合法性將面臨挑戰；反之，隱形禁止具有可適用的合法空間。我們將結合中國法律的既有規定逐一分析。

首先，中國法律法規鼓勵服務提供者積極呈現符合主流價值導向的信息。在對於隱形禁止等利用算法的新型內容審查義務中，值得注意的是《互聯網信息服務算法推薦管理規定》的規定，其在第二章「資訊服務規範」中強調「算法推薦服務提供者不得利用演算法屏蔽資訊、過度推薦、操縱榜單或檢索結果排序、控制熱搜或精選等干預資訊呈現」。[3]但是與此相對的，是該規定的第11條允許並鼓勵算法推薦服務提供者「加強算法推薦服務版面頁面生態管理，建立完善人工干預和用戶自主選擇機制，在首頁、首屏、熱搜、精選、榜單類、彈窗等重點環節積極呈現符合主流價值導向的信息」。相似地，《生成式人工智能服務管理暫行辦法》第5條也「鼓勵生成式人工智能技術在各行業、各領域的創新應用，生成積極健康、向上向善的優質內容，探索優化應用場景，構建應用生態體系」。根據上述正反兩方面的規定，我們認為其存在矛盾之處：一方面，根據文義解釋，為了呈現符合主流價值導向信息的目的，平台有權採取「隱形禁止」措施實施人工干預機制，進行版面頁面的生態管理；但另一方面，平台又具有不得利用算法屏蔽信息、過度推薦、操縱榜單或者檢索結果排序、控制熱搜或者精選等干預信息呈現的義務。兩者如何協調、在實踐中呈現甚麼現狀是值得進一步研究的。

其次，關於隱形禁止措施的合法性問題還需要關注的另一個問題是法律是否規定了平台採取措施後的告知義務。在平台採取內容管理措施的「告知義務」上，歐盟的《數字服務法》第17條的規定可以作為重要參考，該條規定平台採取的每項內容管理措施都應附帶一份「理由說明」（statement of reasons），以提供給受影響的用戶。這一條的規定也被視為研究隱形禁止措施的重要切入點。反觀中國，本文試圖在既有的規範中找到類似對平台採取措施透明性和告知義務的相關條款，但是很可惜，並沒有找到對平台內容審查措施的限制條款。這一現象已被中國多位學者提及，比如梅夏英和楊曉娜（2017）指出，「現有法律的關注重點仍在合理發現義務上，對於技術的獨立性和支配力造成的負面影響沒有納入考慮的範圍中」；張新平（2023）認為，網絡平台治理侵

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

權和失範的風險在增加，但是中國現有的網絡平台治理法律法規對此鮮有規範。由此可以得出初步的結論，中國與其他國家的規定不同，平台在採取內容審核措施時沒有告知受影響用戶的義務，這進一步為平台採取「隱形禁止」措施提供了寬鬆環境和有利條件。

中國平台內容審查的實證研究

這一部分的研究分為兩個部分，第一部分是通過文本分析四大平台的用戶協議，研究其內容審查的標準與舉措；第二部分通過深度訪談和問卷調查，旨在對「隱形禁止」實踐進行剖析。兩部分的研究結果表明，平台內容審查的私權力強化了對用戶表達的審查，容易侵害使用者的合法權利。

平台對在線非法內容審查採取過度謹慎的策略

I. 設定嚴於法律法規的平台非法內容審查標準

研究中，本文分別選取了社交娛樂類和信息資訊類這兩類平台中最熱門的四款應用作為研究對象：微信、抖音、微博、小紅書。這四款應用分別屬於騰訊公司、字節跳動公司、新浪公司、小紅書公司，避免了對同一公司政策進行重覆研究。分析上述四款應用的用戶協議中關於用戶發佈內容審查的標準，可以概括平台對非法內容的標準規定有兩種形式。其中一種形式與《網絡信息內容生態治理規定》類似，試圖儘量詳盡地列舉非法信息的類型。這種形式基本上都是複製《網絡信息內容生態治理規定》中關於「違法信息」和「不良信息」的類型，且不區分兩種非法信息的類型（比如《騰訊微信軟件許可及服務協議》第8.1.2.1條[4]和《「抖音」用戶服務協議》第4.2和4.3條[5]）。此外，平台還會就各自的運營現實適當增加法律法規之外的其他非法信息類型，且標準高於法律法規的要求。比如《騰訊微信軟件許可及服務協議》中新增了四項非法信息的類型（比如使用誇張標題；內容與標題嚴重不符的；炒作緋聞、醜聞、劣跡等的）。另一種形式是直接援引法律法規中的兜底條款，比如《微博服務使用協議》第4.10.1條規定「不得違反中華

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

人民共和國法律法規及相關國際條約或規則」，[6]《小紅書用戶服務協議》第3.2條規定「不得利用小紅書平台製作、複製、發佈、傳播、儲存任何違反法律法規、危害國家安全、可能對用戶財產安全或信息安全造成損失的信息、營銷信息、低俗不當信息、不良信息以及不實信息等內容」。[7]

值得注意的是，各個平台對非法信息的標準設定並不統一。就選取的四款熱門應用而言，微博在其用戶協議中的標準是最為嚴格的，使用「不得違反中華人民共和國法律法規及相關國際條約或規則」這樣的表述將所有法律法規的規定都包括在內。此外，還增加了「不得侵犯任何個人、企業事業單位或社會團體的合法權益」、「不得違反與網絡服務、微博服務有關的網絡協議、規定、程序及行業規則」、「不得以任何方式損害各級國家機關及政府形象」、「不得以任何方式損害微博運營方及其關聯公司的商譽或信譽等合法權益」絕對性的禁止規定。對於法律法規之外新增的「非法信息」標準，各個平台的自由裁量空間很大，平台都沒有給出合理的解釋和相應的法律依據。比如《「抖音」用戶服務協議》曾規定，用戶同意並承諾不製作、複製、發佈、傳播「（17）散佈商業廣告，或類似的商業招攬信息、過度營銷信息及垃圾信息」。商業言論是一項關於經濟利益的表達，維護包含營利目的的廣告性言論自由是經營者在市場經濟中開展各項業務的基礎（龍俊，2023）。中國《憲法》保護言論自由，根據《廣告法》相關規定，對商業言論自由的限制和禁止僅限於特定行業（比如煙草、藥品、轉基因食品等），中國法律法規並沒有對商業廣告的普遍限制。因而《「抖音」用戶服務協議》中「散佈商業廣告」的違法性何在是缺乏依據的。值得注意的是，2024年3月18日更新的《「抖音」用戶服務協議》最新版，在4.2條第（11）款已將該內容修改為「違反網路廣告相關法律法規的商業廣告」。[8]但是《「抖音」用戶服務協議》沒有提供修改前後對照版本，也沒有為用戶提供協議修改的理由。

II. 非法信息處置措施嚴於法律規定

通過文本分析上述四款熱門應用用戶協議中對非法信息內容處置的條款，本文總結平台現有的處置措施類型呈現多樣化，有些措施較之法律規定更為嚴格。上文歸納了當前法律規定了平台處置非法信息的五種

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

主要形式(停止傳輸、防止擴散、消除信息、保存記錄、報告主管部門),也指出了針對算法等新興技術的發展對平台提出的其他處置義務。在四款熱門應用的用戶協議中,處置義務經常體現為禁言、屏蔽信息、刪除發佈內容、封禁賬號、注銷賬號等常見手段。有些平台還發展了新的處置措施,比如《微博服務使用協議》規定平台可以「凍結用戶賬戶資金,用於彌補用戶給微博運營方及其關聯公司、他人造成的損失」。其中,「屏蔽」措施在三個平台(微信、微博、小紅書)中都被明確提及,且沒有限制「屏蔽」措施為用戶可知還是用戶不可知。如果是用戶不可知的屏蔽信息措施,那麼就屬於「隱形禁止」的範圍。

另外,上文規範分析部分指出了現存法律的一個關鍵遺漏,即沒有規定平台如何發現非法信息的程序與約束,而是以平台已經發現非法信息為前提。考察實踐中各大平台的做法,本文發現平台在這一問題上呈現獨斷性,且沒有任何一個平台對自己的自由裁量權做出限制,沒有明確指出用戶提供申訴的渠道。比如《騰訊微信軟件許可及服務協議》規定「騰訊有權不經通知隨時對相關內容進行刪除、屏蔽」,騰訊使用的詞語是「不經通知」、「隨時」。與此類似,抖音和小紅書平台都使用「公司有權獨立判斷」、「採取措施」的表述,同時強調「對於由平台採取措施而造成用戶無法正常使用賬號及相關服務、無法正常獲取賬號內資產或其他權益等後果,公司不承擔任何責任」。

隱形禁止作為平台進行內容審查的重要手段

儘管在廣義上隱形禁止中對「可見性降低」的證明存在難點,但是對秘密的封禁行為的隱形禁止實證調查研究已經存在。事實上,通過判斷用戶對封禁行為是否知情,是了解隱形禁止舉措的一個關鍵切入點。因而我們的研究問題就簡化為:用戶是否經歷過未經通知的封禁?這一部分,本文借鑒了民主和技術研究中心(The Center for Democracy and Technology)2022年4月發佈的一份關於隱形禁止的調研報告(Nicholas, 2022)。研究採用問卷調查和深度訪談兩種研究方法以探討隱形禁止的實例,並對隱形禁止的內容類型、影響效果、救濟途徑等進行研究。值得指出的是,由於隱形禁止本質上幾乎無法被用戶確認,故本研究僅能展開相對主觀的調研,而無法實際證明被「限制可見性」的現實。

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

I. 問卷調查

隱形禁止的證明難點在於，當一個內容沒有像本應該那樣傳播得廣泛時，並不會留下任何痕跡。因此多年來，雖然事實上平台正在利用這些的工具來塑造信息的流動，但是這些方式是不透明的而且難以被證明的。由此，Nicholas（2022）指出，關於這些技術手段最好的證據仍然來自於懷疑自己被隱形禁止的用戶的調查。2023年7月，本研究對設計好的調查問卷進行了小範圍（共計30人）的預調查，並根據預調查的結果對正式調查問卷進行了修改完善。2023年8月至10月，本研究通過問卷星平台提供的樣本服務，發放並回收有效調查問卷362份。通過問卷調查，旨在對普通用戶隱形禁止等經歷進行了解，再結合與網紅用戶、平台工作人員等的深度訪談，希望可以客觀全面地呈現中國平台內容審查的現實。本研究的調查對象來自中國大陸（除新疆地區以外）的各個省份，該樣本呈現多樣化、多元化，符合問卷調查的覆蓋性要求（詳見表一調查問卷樣本分析）。

表一　調查問卷樣本分析

性別比例	每天上網時長	年齡段
	1小時以下（2.75%）	18歲以下（0.83%）
男（48.76%）	1–3小時（26.45%）	18–24歲（18.18%）
	3–6小時（37.47%）	25–34歲（57.02%）
	6小時以上（31.40%）	35–44歲（15.98%）
女（51.24%）	不上網（1.93%）	45–54歲（4.41%）
	／	55歲以上（3.58%）

調研結果分析如下：

第一，普通用戶較少經歷傳統的平台內容審查（比如禁言、封號等），僅有33.33%的用戶表示有被平台禁言、屏蔽或封禁賬號的經歷。這表明，事實上對於普通的用戶，發佈內容多屬於「安全範疇」，遭到平台內容管理的比例並不高。在那些選擇了被禁言、屏蔽等經歷的用戶中，82.64%的用戶表示平台採取舉措是有通知的。但是對於被禁言的原因，仍有約52.89%的人選擇了「不知道」和「比較模糊，不清楚」。由此可知，即使在用戶被告知的情況下，用戶仍對平台內容管理的規則和理由欠缺理解，這表明平台發佈的「告知內容」存在問題。

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

第二，與傳統封禁的事實經歷不同，用戶對於平台實施「隱形禁止」有更多共識。在問卷的第三部分，關於「您覺得平台是否有對您秘密地實施過禁言、降低內容可見度等舉措」的回答中，選擇「肯定有」(27%) 和「應該有」(41.60%) 的用戶佔到了68.60%。雖然78.24%的受訪者沒聽說過「隱形禁止」術語或不理解該術語的含義，但是受訪者卻會勾選以下經歷 (隱形禁止比較典型的表現)：感覺自己的參與度被降低了 (比如獲得更少的點讚、評論、分享、閱讀等) (51.24%)；發佈的內容沒有出現在某些特定的標籤頁面或話題之下 (33.33%)；從其他人的賬號處看不到自己的首頁、狀態和評論等 (31.13%)；更新的狀態沒有推送給自己的粉絲或關注自己的人 (25.90%)；對方看不到自己在聊天框裏看到的信息 (23.97%)。換言之，抽象的隱形禁止概念相對於普通用戶而言是比較陌生的，但是具體的隱形禁止舉措和經歷卻是很常見的。此外，受訪者結合自身經歷，認為最容易出現過隱形禁止行為的平台類型排序為：內容共享平台 (比如微博、播客)、社交聊天平台 (比如微信、QQ)、短視頻平台 (比如抖音、快手)、問答推薦平台 (比如小紅書、知乎、豆瓣)、電商平台 (比如淘寶、天貓)、外賣平台 (比如美團、餓了麼)，以及生活娛樂推薦平台 (比如大眾點評)。

第三，關於被內容審查後的應對舉措，近一半 (52.07%) 受禁言的用戶選擇找平台申訴，但是從申訴結果來看，只有30.58%的用戶得到了有效的解決。另一半沒有申訴的用戶，或是因為找不到申訴渠道 (27.27%)，或是「躺平」等待自然解禁 (20.66%)。可見平台和用戶之間申訴解決機制並不完善，用戶維權動力不足。

通過問卷調查，我們了解了普通用戶對平台內容管理的整體感知，通過受訪者對「禁言後是否對您的工作生活造成影響」問題的回答，我們可以看到，33.06%的用戶表示「沒有影響」，「影響很大」的用戶僅為5.79%。於是我們將進一步通過深度訪談，聚焦特定更容易受到平台內容管理的活躍用戶，對「隱形禁止」展開更為深入的研究。

II. 深度訪談

本研究在深度訪談部分，區分用戶方和平台方作為兩類訪談對象。前者為用戶側，包括普通個人用戶、網紅用戶 (粉絲量大於六千)、企業

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

用戶，也包括幫助運營賬號的營銷公司等；後者包括平台負責內容審查的工作人員、技術人員等。從2023年8月至11月，共20位受訪者接受了深度訪談，其中平台方和用戶方各佔一半，18名受訪者接受了電話訪談，時間不少於20分鐘；另有兩名受訪者根據問題列表提供了書面答覆。在研究中，由於內容審查這一話題的敏感性，多位聯繫到的受訪者（不論是平台方還是用戶方）都表示不願意接受訪談，即使匿名訪談也被拒絕。這一問題在國外的研究中也有被提及，比如Gillespie（2022）在其研究中表示，正式的研究訪談需要和訪談對象簽訂保密協議（non-disclosure agreement, NDA），但是保密協議會成為公開談話內容的實質性障礙。而如果是非正式的訪談，受訪者又會考慮到研究人員背後複雜的商業和工作關係而有所保留，這些是研究中遇到的方法論問題。而這一問題在中國語境下則更為嚴峻。表二列出了訪談對象的構成，他們橫跨了短視頻、直播、電商、生活娛樂、社交等熱門類型的應用，覆蓋了中國最為重要的幾大互聯網平台，具有典型性和代表性。

表二　訪談對象結構

訪談類型	編號	職業
A用戶方	A1	普通用戶
	A2	企業用戶
	A3	網紅用戶
	A4	網紅用戶
	A5	網紅用戶
	A6	內容運營機構
	A7	網紅用戶
	A8	企業用戶
	A9	普通用戶
	A10	普通用戶
B平台方	B1	平台內容審核員
	B2	平台內容審核員
	B3	平台內容審核員
	B4	算法工程師
	B5	算法工程師
	B6	平台內容審核員
	B7	平台內容安全負責人
	B8	軟件工程師
	B9	平台內容審核員
	B10	企業法務

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

通過深度訪談，本研究有以下三點發現：

第一，受訪者對「隱形禁止」的普遍存在達成共識，並列舉了一些典型的隱形禁止情形。對於平台秘密地限制內容可見性的做法，受訪者都給予了肯定的回答，多數受訪者直接表示「隱形禁止在實踐中非常常見」。同時，不同平台和用戶也提供了一些比較典型的「隱形禁止」做法，比如在評論區設置「僅用戶可見」(B1)、對問題內容呈現在搜索結果的前十位以外(B5)、內容可以正常發佈但是瀏覽量為零(A7)等。各個平台根據自身的性質和商業模式，有不同的「隱形禁止」手段。比如在直播平台中，利用發佈者和接收者的短暫幾秒的時間差，平台可以過濾、消音平台認為不合適的內容。用戶A1提到企業公眾號發佈關於並不受推崇的「萬聖節」相關視頻案例中，前一週視頻獲得了超級高的流量，但是到了第二週，流量突然呈現斷崖式下跌：「很明顯，平台對相關視頻進行了限流操作」。與此同時，平台方的受訪者也從自己的工作經歷中，豐富了實踐中平台採用隱形禁止舉措背後的原因。從隱形禁止的驅動力來看，除了平台合規義務的外動力，也有平台自身為營造良性平台生態環境的內動力：「在類似的短視頻平台上熱播的內容，在我們平台未必會被推薦，因為我們不希望把我們的平台弄得『烏煙瘴氣』」(B3)。此外，平台方還提出，不公開這些具有操作性的內部標準，是防止用戶「鑽空子」，反噬利用規則(B7)。

第二，雖然受訪者表示隱形禁止的做法很普遍，但是用戶多認為完全的「隱形」是不太現實的，因為他們雖然沒有被告知，但是基本上他們可以很快知道自己被「限流」。A7表示自己曾發佈一張露背的照片，結果幾分鐘過去了，發現瀏覽量為零，她覺得這就是典型的「隱形禁止」:「所以基本上用户可以知道自己被限流，只是可能並非第一時間知道」(A7)。同時，作為用戶，他們每個人都在避免內容審核上探索和發展出了豐富的「民間理論」(folk theories)。所謂「民間理論」，Moran等人(2022)指出，是一種用戶培養出的關於平台可見性如何運作的專業知識策略。A5是一位粉絲量超二百萬的用戶，他得意地總結自己的成功經驗有三點：第一，個人涉足互聯網內容領域時間長，資歷深，非常了解規則；第二，個人與平台、政府都有良好的互動關係；第三，職業敏感性強，只討論專業的技術問題，絕不涉及政治議題的討論。A3則認為發佈內容的首圖、發佈時間是比較重要的，同時避免對太新的事件

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

發表看法。A6表示自己專門幫忙推廣企業賬號，成功的秘訣就是「通過多年來的不斷試錯，我們比一般人更了解規則」。

第三，受訪者的另一個普遍共識是平台在內容管理上享有很大的權力，可以通過四個方面反映出來。首先，平台內容管理的規則不透明，自由裁量權很大。目前各大互聯網平台對內容審查都有規範化的流程和標準，一般包括算法初審、人工審查、專業人員覆審三個步驟，每個平台都有專門培訓內容審查的工作人員。其中，內容審查的標準是被公認為最具有挑戰性的問題，因為法律法規的規定是比較模糊和原則性的，如何將合規義務體現為具體標準，每個平台都在摸索中，也都形成了自己的內容審查標準。一般來說，審查標準分為對內和對外兩種，對內的培訓手冊會明確列舉內容審查的「關鍵詞」文本，這些「關鍵詞」標準是不公開的，僅供平台內部人員使用，而且更新頻率非常快；甚至在不同時間節點，審查的關鍵詞和內容也存在不同。對外的內容標準一般是倡導性的指南和規範，用以引導用戶發佈健康積極的內容。比如「低胸」照片被視為容易引發低俗內容的「不良內容」，在對「低胸」的具體判斷標準，平台內容有非常細節的規定，包括具體位置、厘米數等（B6）。雖然經過專門培訓的審查人員和算法工程師非常熟悉內容審查的「關鍵詞」，但是這些詳細的「關鍵詞」標準是由誰制定、如何制定等，他們並不清楚。

其次，平台有些審查內容的規則（包括非法內容標準及採取措施標準兩方面）是令人費解的，這些影響了使用者正常的表達自由。A6提到，視頻中不能出現「腎」這個字，這一點「以偏概全」，但是平台就是這樣規定的。作為平台內部的工作人員，有時候也不清楚規則。A4談到之前在某平台工作的經歷，平台自己的賬號由於贛南地區地圖上的小錯誤而未被審核通過，「如果不是在平台工作，一般用户根本無法獲知被禁言的理由」。類似地，B3也曾幫朋友詢問一則視頻被屏蔽的原因，是因為背景音樂中含有特殊少數民族的語言，「這一點作為平台內容審核人員，我們自己也很驚訝」。從平台採取的措施來看，平台審查內容的選項區分為全部可見、部分可見、僅自己可見、刪除，不同具體選項的背後帶有不同的權力範圍和影響，但是沒有哪一個平台對此有具體說明和解釋。

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

再次，在內容管理的申訴上，雖然平台強調其申訴渠道24小時暢通，但是用戶普遍表述申訴的無用性。B7表述，根據工信部、網信辦等部委要求，用戶可以通過客服電話或是現場客戶接待窗口申訴對應問題。但是A3用戶指出，所謂的平台規則都是「假大空的話」，真正的解釋權和決定權全在平台。A2結合自身被禁言的經歷表示，平台對於申訴只會給予程式化的回覆，即簡單列舉規則的條款「你的內容違法了用戶協議第X條」。即使他們事後還專門找到工作人員了解情況，得到的回覆仍然是「算法打擊無誤，但是不方便透露具體的禁言原因」。

最後，一些用戶甚至擔心被平台監聽而不願意接受我們的非正式訪談。A3提醒到，一些用戶擔心自己接受訪談的行為會被認為是詆毀平台的行為：「有些博主甚至不給自己的狀態點讚，因為擔心平台看到內容發佈者和點讚者同屬一個IP地址，而認定內容違法」。事實上，對於明顯違法的內容，一般不用太過擔心，因為會有很多用戶舉報違法內容。但是對於涉及商業利益的內容排序，標準就很模糊了，平台的自主權很大，用戶沒有表達不滿的途徑和反抗平台權力的能力。由此，在用戶中流傳著一種說法，內容流量是一門「玄學」(A2)。平台方表示流量分配是現實中的一個難題(B7)。B5從技術角度分析，「玄學」的說法是有一定道理的，因為內容呈現是深度學習和人工參與的組合。

此外，受訪者還提到了由平台私權力過大引發的另一個現實風險，即平台「權力尋租」問題已經比較普遍了。具體來說，平台內部工作人員對於每個領域都有自主決定權，分配一些加碼的流量，賄賂工作人員以獲得更多流量推薦的情況在現實中比較常見(A3)。一些美妝博主和平台審核人員關係好，博主贈送禮物給工作人員，工作人員分配流量給博主，這是一種「雙贏」(A7)。有些平台是有明確的「商業化中心」，對於花多少錢來增加粉絲量和增加閱讀量是有明碼標價的(A1)。這也是類似於A6這樣「運營賬號」的中間商工作的主要內容。

結論與建議

當今數字時代，「信息是權力的中心」，信息技術的發展使得信息呈現中樞化和權力化趨勢(馬長山，2022)，平台獨特的優勢和地位讓

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

他們獲得了強大的「數據權力」（尹華容、王惠民，2022）。對於平台封禁行為而言，它已不再是平等主體之間的合同關係所能闡述和包含的，其實質是平台經濟業態塑造了數字時代的信息生態，而這種生態顛覆了傳統生活方式和行為模式，變成了進入世界、觀察世界、參與世界的全新基本途徑（馬長山，2022）。上文從規範分析和實證分析兩個維度，結合中國現實試圖對「隱形禁止」這一複雜問題做出探討，得出「平台內容審查存在失範風險」的結論，下文將對此提出約束平台內容審查權力的針對性建議。

平台內容審查存在失範風險

I. 規範分析：欠缺對平台權力的約束條款

隨著數字時代平台經濟的縱深發展，鑒於平台在數字經濟中的支配力和影響力，在對在線非法內容的審查中，中國法律和各國法律類似，都重視充分發揮平台的私權力，彌補政府規制能力的不足。平台的主體責任已經出現在大量規範性文件中（比如《關於進一步壓實網站平台信息內容管理主體責任的意見》、《互聯網平台落實主體責任指南（徵求意見稿）》等），也日益廣泛地出現在法律法規等規範中（劉權，2022）。具體對在線內容的審查中，相關規範更是直接明確「網絡信息內容服務平台應當履行信息內容管理主體責任」。[9] 在線內容審查的立法核心是強調平台的主體責任，始終圍繞平台的責任和義務具體展開條文的設計。但是當前中國法律對規範平台自治權的立法不足，缺乏對平台治理在線非法內容失範風險的約束。具體來說，現階段法律法規對網絡平台的授權仍處於粗放狀態，既沒有對平台在線內容審查的權力授予機制做出明確的規定，也沒有對平台實施權力行為的程序做出約束，放大了網絡平台權力運行的失範風險（郭漸強、陳榮昌，2019）。

典型的表現是忽視用戶權利保障機制的規定。網絡平台權責結構設計存在一定差序化傾向，權責對應性框架尚未形成，主要體現在重視網絡平台的國家義務設定，用戶權利保護責任設計則相對被邊緣化（郭漸強、陳榮昌，2019）。對於平台用戶而言，面對平台的治理優勢

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

和技術效能，平台的自治權逐漸壯大為「政治權力與市場權利之外的第三種力量」(李曉宇，2022)，無形中會對用戶的自由選擇權、知情權、公平交易權、表達權帶來挑戰(陳兵，2020)。研究發現，當前法律法規對用戶權利保護的條款多為宣示性條款，沒有對平台內容審查規則的透明度提出義務，沒有規定公開透明的平台權力行使程序，也缺乏對用戶權利救濟機制(比如申訴維權等)的義務要求。有趣的是，《網絡信息內容生態治理規定》中雖然規定了平台受理用戶的投訴、舉報制度，但是制度設計是針對網上違法和不良信息，旨在鼓勵網絡信息內容服務使用者積極參與網絡信息內容生態治理，[10] 並不是針對平台單方面實施的刪帖、封號、禁言等行為的申訴維權渠道。

除此之外，「隱形禁止」影響意見的能力也是令人擔憂的，其最常被指摘的危險在於平台可以透過選擇目標使用者來接收特定內容實現操縱觀點、影響意見，從而被用作政治宣傳的手段，例如Twitter和Facebook被指控秘密地透過隱形禁止壓制保守派線上內容發佈(Chen & Zaman, 2024)。與此類似，劍橋分析數據醜聞中曾採用「隱密帖」(dark post)的形式進行宣戰宣傳，進而影響美國大選。

II. 實證分析：平台內容審查私權力已呈現侵蝕用戶權利的現實風險

實證研究表明，平台在內容審查的私權力強化了對用戶表達的審查，平台過分追逐商業利益而未兼顧公平、侵蝕用戶權利的風險已經突顯。首先，平台為安全地規避合規風險，往往傾向於選擇高於法律標準的內容審查標準和舉措，隱形禁止等隱蔽的審查更容易對使用者表達自由造成侵害。加之當前法律缺乏約束平台權力的規定和用戶維權的渠道，平台通過內容審查侵犯用戶權益的風險大大加強。研究表明，這種風險對依賴平台作為經濟收入來源的「網紅」用戶的影響尤其明顯。這種風險不僅會對用戶個人產生深刻影響，長此以往，普通用戶屈服於平台的私權力之下，必定會對整體數字市場秩序造成深遠影響，最終將損害社會的整體利益。

其次，隱形禁止在平台內容審查的實踐中是常見的手段，為其追逐商業利益提供了便利條件。平台選擇隱形禁止的動機是多元的，可能是因為防止用戶逃避規則，可能是因為隱形禁止的溫和性，也可能

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

是因為內容審查規則本身的複雜性和不易公開性。在中國，常見的隱形禁止措施體現為用戶發佈的內容僅自己可見、發佈的內容沒有出現在某些特定的標籤頁面或話題之下、對方看不到用戶在聊天框裏發佈的信息等。超級的數字能力和優勢下，平台通過隱形禁止等手段，更有能力也更有動力對用戶呈現的內容進行管理、以及決定呈現甚麼給用戶看以控制用戶的判斷和影響。

再次，用戶現實維權陷入了困境，發起救濟的私主體收益與能力不足，缺乏維權動機。在面對平台過度強大的內容審查私權力時，個人發起維權的能力較弱，即使具有了收集證據等能力，維權結果也是行政處罰，對於消費者的維權動力不足（張凌寒，2018）。這事實上進一步加強了平台在內容控制上的權力，與最初設置平台責任的初衷背道而馳。當前這種不斷集中的強大權力已經引發了公眾的強烈擔憂，也成為各國立法者與司法裁判者努力破解的一道法律難題。

約束平台內容審查權力的建議

面對「隱形禁止」等新興內容審查方式，結合上文對中國既有法律的規範分析和實踐研究，本文提出三點規範平台內容審查的建議。

I. 以用戶協議為抓手，明確平台內容審查的規則

從性質來看，《用戶協議》是平台與用戶之間的合同，雙方都有權依據合約約定主張協議的內容。但平台在技術、經濟實力等方面具有顯著的優勢，很容易利用這些優勢制定不合理的《用戶協議》，擴大自己在內容審查中的權利。上文已詳細列舉了平台在《用戶協議》中擴大自身權利、限制使用者合法權益的條款，這些都侵害到了使用者表達自由、合理申訴等權利。因而可以《用戶協議》為規範平台內容審查實務的重要抓手，透過協議的制定、發佈、變更等程序，確認其法律效力，更好地平衡用戶與平台之間的關係。這種想法的轉變也被認為是從外在問責為導向的監管，轉向以審查格式合約和網規為重點的監管，旨在加強用戶協議的透明性和有效性，幫助用戶加強介入協議談判的理論（胡凌，2019）。《用戶協議》的制定和修改應充分吸收使用者

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

的合理建議，保證協議條款的公開透明，並增強平台內容審查規則的可預見性。尤其在內容審查事項中，要防止平台制定免除其責任、加重對方責任、排除對方主要權利的條款，確保協議的合法性。同時《用戶協議》的變更在性質上屬於合約變更，其應及時向用戶說明且取得用戶的同意，否則對用戶不發生法律效力。此外，協議中也應注重對用戶申訴、維權等措施的規定，調節用戶與平台在內容審查上的力量懸殊，以更好地促進彼此的信任，共建和諧可信的平台生態。

II. 以算法技術為切入，規範平台內容審查的實踐

目前，算法推薦內嵌於內容分發機制，為平台提供了全新的內容可見度與注意力分配機制。人們懷著對技術「客觀中立」的想像，容易忽略了演算法實現過程中的主觀性。反映在平台內容審查中，突出的特點是內容可見度與平台商業利益緊密捆綁，非中立性和非自動性特徵明顯。雖然從2017年的《電子商務法》開始，中國法律就初步建立了演算法事前監管制度，在事前對搜尋演算法的多樣性、客觀性提出了監管要求，且明確了網絡交易平台因不當部署、應用算法而承擔的法律責任。但是面對隱形禁止等新興內容審查方式，算法技術在平台內容審查中發揮的作用越來越重要，平台內容審查權責邊界的明確可以算法技術的使用和規範為切入點，從算法運行的正當程序和問責機制等面向約束平台內容審查的權力。一方面，加強平台對算法使用的自律機制，包括定期公開內容審查算法運行報告、提交算法備案、落實算法審計制度等，降低部署算法應用中的異化風險。另一方面，配置個人權利以對抗算法權力，例如賦予個人獲得算法推薦內容的解釋說明權、允許個人退出算法決策並尋求人工決策的權利、落實個人不受純粹算法決策支配的權利等。

III. 以公開透明為目標，監督平台內容審查的權力

平台內容審查權力失範的另一個重要原因在於對平台內容審查實務缺少程序上的約束。針對算法和人工智慧技術應用中存在的問題，中國近年來出台了多項法規，但是值得注意的是，現有規定多為原則性、整體性規範，在針對「隱形禁止」等平台內容審查的新興措施顯得

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

較為乏力。現有規則缺乏對內容審查公開性和透明性的約束，沒有對平台內容審查程序正當原則的要求，這就導致了實踐中各大平台「有權不經通知隨時對相關內容進行刪除、屏蔽」。當使用者在現實中遇到平台單方面實施的禁言、封號等內容審查措施時，缺乏必要的申訴維權管道，應從法律層面對平台權力行使的限制和用戶維權管道做出明確規定。就此問題，歐盟2022年通過的《數字服務法》從透明度和程序正當角度對平台進行內容管理做出規範，可以成為中國約束平台內容管理權力的有益範本。《數字服務法》對於在線內容管理提出了一套全面的專門規範，核心範疇是「數字守門人」被要求承擔的諸多平台責任。其核心框架可以概括為「通知與採取行動」（notice-and-action），即平台在進行內容管理時，應遵循程序正當原則，包含三個步驟：第一，每個制裁的措施必須受到明確和可預見的約束；第二，制裁措施應該向受到影響的用戶進行通知和解釋；第三，受影響用戶必須被允許提出上訴和反對。其中，法案第17條細化了平台採取措施需要說明的內容，即要求平台對受到影響的用戶提供說明。《數字服務法》專門規定了推薦系統的透明度要求，即在線平台應當在其用戶協議中以通俗易懂的語言列出主要參數，解釋清楚為甚麼要想服務接收者推薦有關的信息內容。如果在線平台提供了多種推薦系統，還應同時設置提供用戶隨時選擇並修改首選項（程寶庫、李俊輝，2023）。這些域外最新的規範可以成為中國完善平台權力約束的有益參考。

註釋

1 《數字服務法》條款第55條原文如下：「Restriction of visibility may consist in demotion in ranking or in recommender systems, as well as in limiting accessibility by one or more recipients of the service or blocking the user from an online community without the user knowing it (“shadow banning”)」。其他「限制可見性」還可能表現為在排名或推薦系統中的降級、限制服務接收者的可訪問性等。

2 《互聯網信息服務管理辦法》第15條的「九不准」如下：（一）反對憲法所確定的基本原則的；（二）危害國家安全、洩露國家秘密、顛覆國家政權、破壞國家統一的；（三）損害國家榮譽和利益的；（四）煽動民族仇恨、民族

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

歧視、破壞民族團結的；(五) 破壞國家宗教政策、宣揚邪教、封建迷信的；(六) 散佈謠言、擾亂社會秩序、破壞社會穩定的；(七) 散佈淫穢、色情、賭博、暴力、兇殺、恐怖或者教唆犯罪的；(八) 侮辱或者誹謗他人，侵害他人合法權益的；(九) 含有法律、行政法規禁止的其他內容的。

3 《互聯網信息服務算法推薦管理規定》第14條：「算法推薦服務提供者不得利用算法虛假註冊賬號、非法交易賬號、操縱用戶賬號或者虛假點讚、評論、轉發，不得利用算法屏蔽信息、過度推薦、操縱榜單或者檢索結果排序、控制熱搜或者精選等干預信息呈現，實施影響網絡輿論或者規避監督管理行為。」

4 深圳市騰訊電腦系統有限公司 (2024)。〈騰訊微信軟件許可及服務協議〉。取自微信，https://weixin.qq.com/cgi-bin/readtemplate?lang=zh_CN&t=weixin_agreement&s=default#。

5 北京抖音科技有限公司 (2024)。〈「抖音」用戶服務協議〉。取自抖音，https://www.douyin.com/draft/douyin_agreement/douyin_agreement_user.html?id=6773906068725565448。

6 〈微博服務使用協議〉(2024)。取自微博，https://weibo.com/signup/v5/protocol。

7 〈小紅書用戶服務協議〉(2024)。取自小紅書，https://www.xiaohongshu.com/crown/community/terms。

8 請見《「抖音」用戶服務協議》第4.2條第 (11) 款。需要說明的是，筆者已找不到舊版的《抖音用戶服務協議》。

9 《網路信息內容生態治理規定》第8條：「網路資訊內容服務平台應履行資訊內容管理主體責任，加強本平台網路資訊內容生態治理，培養積極健康、向上向善的網路文化。」

10 《網絡信息內容生態治理規定》第16條：「網絡信息內容服務平台應當在顯著位置設置便捷的投訴舉報入口，公佈投訴舉報方式，及時受理處置公眾投訴舉報並反饋處理結果。」第20條鼓勵網絡信息內容服務使用者積極參與網絡信息內容生態治理，通過投訴、舉報等方式對網上違法和不良信息進行監督，共同維護良好網絡生態。為此，中國網信辦也專門推出了「違法和不良信息舉報平台」(https://www.12377.cn/)。

披露聲明

本文作者未報告潛在的利益衝突。

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

Disclosure Statement

No potential conflict of interest was reported by the author.

研究經費資助

本文受教育部人文社會科學研究青年計劃「生成式人工智慧背景下網路內容治理的平台權責邊界研究」（24YJC860033）、中國政法大學科學研究創新年度青年計劃「智慧傳播中資料要素應用的風險與防範研究」（24KYQN022）資助。

Funding

This research was funded by Youth Project of the Ministry of Education for Humanities and Social Sciences Research, "Research on the Platform's Rights and Responsibilities Boundary in Online Content Governance in the Context of Generative Artificial Intelligence" (24YJC860033) and Annual Youth Research Innovation Project of China University of Political Science and Law, "Research on Risks and Prevention of Data Element Application in Intelligent Communication" (24KYQN022).

ORCID

張璐（Lu ZHANG）https://orcid.org/0000-0003-2544-7792

參考文獻

中文部分（Chinese Section）

孔祥穩（2020）。〈網絡平台資訊內容規制結構的公法反思〉。《環球法律評論》，第2期，頁133–148。

Kong, X. (2020). A public law reflection on the regulatory structure of information content on online platforms. *Global Law Review*, *2*, 133–148.

尹華容、王惠民（2022）。〈數據權力的興起、異化及規制〉。《湖南大學學報（社會科學版）》，第3期，頁124–132。

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

Yin, H., & Wang, H. (2022). The rise, alienation, and regulation of data power. *Journal of Hunan University (Social Sciences)*, *3*, 124–132.

李曉宇(2022)。〈「元宇宙」下虛擬數據作品的著作權擴張及限制〉。《法治研究》,第2期,頁15–24。

Li, X. (2022). Copyright expansion and limitations of virtual data works in the "metaverse." *Legal Research*, *2*, 15–24.

姚志偉(2019)。〈技術性審查:網絡服務提供者公法審查義務困境之破解〉。《法商研究》,第1期,頁31–42。

Yao, Z. (2019). Technical review: Breaking the dilemma of public law review obligations for online service providers. *Studies in Law and Business*, *1,* 31–42.

胡凌(2019)。〈從開放資源到基礎服務:平台監管的新視角〉。《學術月刊》,第2期,頁96–108。

Hu, L. (2019). From open resources to basic services: A new viewpoint of platform supervision. *Academic Monthly*, *2*, 96–108.

馬長山(2022)。〈數位法理的理論表達〉。《中國法學》,第3期,頁119–144。

Ma, C. (2022). The theoretical expression of digital jurisprudence. *China Legal Science*, *3*, 119–144.

張凌寒(2018)。〈《電子商務法》中的算法責任及其完善〉。《北京航空航太大學學報(社會科學版)》,第6期,頁16–21。

Zhang, L. (2018). Algorithmic liability in *Electronic Commerce Law* and its improvement. *Journal of Beijing University of Aeronautics and Astronautics (Social Science Edition)*, *6*, 16–21.

張晨穎(2021)。〈公共性視角下的互聯網平台反壟斷規制〉。《法學研究》,第4期,頁149–170。

Zhang, C. (2021). Anti-monopoly regulation of Internet platforms from the perspective of publicness. *Chinese Journal of Law*, *4*, 149–170.

張新平(2023)。〈網絡平台治理立法的反思與完善〉。《中國法學》,第3期,頁122–141。

Zhang, X. (2023). Reflections and improvements on the legislation of online platform governance. *China Legal Science*, *3*, 122–141.

梅夏英、楊曉娜(2017)。〈自媒體平台網絡權力的形成與規範路徑〉。《河北法學》,第1期,頁36–47。

Mei, X., & Yang, X. (2017). The formation and regulatory path of network power in self-media. *Hebei Law Studies*, *1*, 36–47.

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

郭漸強、陳榮昌（2019）。〈網絡平台權力治理：法治困境與現實出路〉。《理論探索》，第4期，頁116–122。

Guo, J., & Chen, R. (2019). Governance of power on online platforms: Legal dilemmas and practical solutions. *Theoretical Exploration*, *4*, 116–122.

陳兵（2020）。〈因應超級平台對反壟斷法規制的挑戰〉。《法學》，第2期，頁103–128。

Chen, B. (2020). Responding to the challenges of antitrust regulation posed by super platforms. *Jurisdiction*, *2*, 103–128.

程寶庫、李俊輝（2023）。〈非法內容審查的平台責任研究——歐盟《數位服務法》的分析與借鏡〉。《中國市場監管研究》，第2期，頁20–25。

Cheng, B., & Li, J. (2023). Research on platform responsibility for illegal content review: Analysis and lessons from the EU's *Digital Services Act*. *Research on China Market Supervision*, *2*, 20–25.

趙鵬（2016）。〈私人審查的界限——論網絡交易平台對用戶內容的行政責任〉。《清華法學》，第6期，頁115–132。

Zhao, P. (2016). The limits of private review: Administrative responsibility of online transaction platforms for user content. *Tsinghua Law Review*, *6*, 115–132.

劉權（2020）。〈網絡平台的公共性及其實現〉。《法學研究》，第2期，頁42–56。

Liu, Q. (2020). Public nature of internet platform and its realization. *Chinese Journal of Law*, *2*, 42–56.

劉權（2022）。〈論互聯網平台的主體責任〉。《華東政法大學學報》，第5期，頁79–93。

Liu, Q. (2022). On the subject responsibility of internet platforms. *Journal of East China University of Political Science and Law, 5*, 79–93.

龍俊（2023）。〈商業言論自由的管制邊界〉。《環球法律評論》，第4期，頁73–89。

Long, J. (2023). The regulatory boundaries of commercial free speech. *Global Law Review*, *4*, 73–89.

英文部分（English Section）

Chen, Y.-S., & Zaman, T. (2024). Shaping opinions in social networks with shadow banning. *PLoS ONE, 19*(3), e0299977.

Cotter, K. (2019). Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram. *New Media & Society, 21*(4), 895–913.

Copyrighted material of: School of Journalism and Communication, The Chinese University of Hong Kong;
School of Communication, Hong Kong Baptist University (2025). Published by The Chinese University of Hong Kong Press.
ALL RIGHTS RESERVED.

Gillespie, T. (2022). Do not recommend? Reduction as a form of content moderation. *Social Media + Society*, *8*(3), 20563051221117552.

Goldman, E. (2021). Content moderation remedies. *Michigan Technology Law Review*, *28*(1), 1–59.

Leerssen, P. (2023). An end to shadow banning? Transparency rights in the *Digital Services Act* between content moderation and curation. *Computer Law & Security Review, 48*, 1–10.

Moran, R. E., Grasso, I., & Koltai, K. (2022). Folk theories of avoiding content moderation: How vaccine-opposed influencers amplify vaccine opposition on Instagram. *Social Media + Society*, *8*(4), 1–9.

Nicholas, G. (2022, April 26). *Shedding light on shadowbanning*. Center for Democracy & Technology. Retrieved from https://cdt.org/insights/shedding-light-on-shadowbanning/.

Radsch, C. (2021). Shadowban/Shadow banning. In L. Belli, et al. (Eds.), *Glossary of platform law and policy terms*. Internet Governance Forum.

Savolainen, L. (2022). The shadow banning controversy: Perceived governance and algorithmic folklore. *Media, Culture & Society*, *44*(6), 1091–1109.

本文引用格式

張璐（2025）。〈中國語境下平台內容審查的權力及其界限——以隱形禁止為例〉。《傳播與社會學刊》，第71期，頁183–209。

Citation of This Article

Zhang, L. (2025). The power and limits of platform content moderation in the Chinese context: Taking shadow banning as an example. *Communication and Society*, *71*, 183–209.