

מיפוי מנגנוני הגנה על בחירות והשלמות נדרשות בשירותים דיגיטליים בישראל: Meta
WhatsApp / Facebook / Instagram / Meta AI & Muse / Threads

מדיניות ומכשירים ייעודיים לבחירות שהופעלו מחוץ לישראל	פעילות ייעודית לזירה הישראלית
מרכזי בחירות ייעודיים שהוקמו להגנת בחירות, לניטור בזמן אמת ושילוב גופים מקומיים: הודו (2024), ברזיל (2014), קנדה (2025), ארה"ב (בחירות האמצע 2026) ועוד. [4] [5] הבלטה והנגשת מידע בחירות: בבחירות בקנדה ב-2025 ובחירות אמצע בארה"ב ב-2026 הופעלו הודעות בראש הפיד של Facebook ו-Instagram, התראות הצבעה והפניות בחיפוש למקורות רשמיים. בארה"ב קיימות התאמות נוספות בהתאמה לשפה ומיקום. [4] [5] שת"פ ייעודי עם ארגונים מקומיים: בחירות כלליות בהודו 2024: קו WhatsApp לבדיקת deepfakes, לצד 12 שותפי בדיקת עובדות ב-16 שפות מקומיות וערוץ בעדיפות גבוהה לוועדת הבחירות של הודו לסימון תוכן בלתי חוקי; ⁶ בחירות מוניציפליות בברזיל 2024: הרחבת שת"פ עם ארגוני בדיקת עובדות מקומיים ורשות הבחירות, לרבות ערוצי דיווח והנגשת מידע רשמי. ⁷ פרסום פוליטי ותוכן סינתטי (מדיניות גלובלית, חלה גם בישראל): אימות מפרסם, הצהרת "Paid for by" ושמירה ב-Ad Library; חובת גילוי של שימוש ב-AI במודעות. תווית AI info ב-Facebook, Instagram ו-Threads באמצעות אותות סימון או גילוי עצמי של המשתמש המעלה. ⁸ אכיפה ייעודית לזיהוי וסיכול רשתות השפעה: אכיפה נגד התערבות בבחירות ושיבוש רשתות Coordinated Inauthentic Behavior; בגרמניה 2025 תועדה היערכות רב-פלטפורמית, לרבות כללי קמפיינים, פרסום ותוכן פוליטי. ⁹ מחקר ושקיפות: Meta Content Library ו-API כוללים תוכן ציבורי מ-Facebook ו-Instagram (ומספטמבר 2024 גם מ-Threads, בממשק המשתמש), וחוקרים ממוסדות אקדמיים או מלכ"רים מתאימים יכולים להגיש בקשה לגישה, בכפוף לבחינה. ¹⁰ WhatsApp: מנגנוני מוצר גלובליים זמינים למגבלות העברה (הודעה שהועברה – לקבוצה אחת בכל פעם) ולצמצום ויראליות; ¹¹ WhatsApp Business Platform אסורה לשימוש מפלגות, פוליטיקאים, מועמדים וקמפיינים. ¹²	לא נמצא תיעוד פומבי למרכז בחירות היערכות ייעודית לבחירות ישראל 2026 באף אחד משירותי החברה. פרסום ממומן: מאז 2019 מופעלים בישראל מנגנון אימות מפרסמים, גילוי זהות המממן ושמירת מודעות פוליטיות ובנושאים חברתיים ב-Meta Ads Library. ¹ אולם הצהרה זו אינה ייעודית לבחירות 2026. כלי אכיפה ומניעה ייעודיים: ב-2022 הושק מסווג עברית לזיהוי יזום ואכיפת תוכן מפר; בנובמבר 2024 הוטמע ניתוב תוכן בערבית לפי ניב ושודרג ערוץ ההסלמה של Trusted Partners. ² ³

¹ Meta, **Expanding Our Efforts to Protect Elections in 2019** (28.1.2019): <https://about.fb.com/news/2019/01/elections-2019/>

² Meta Final Update: **Israel and Palestine Human Rights Due Diligence** (Dec. 2025), including Arabic-dialect routing and Trusted Partner escalation updates; the report also clarifies that Facebook/Instagram content-moderation commitments do not automatically apply to WhatsApp: <https://humanrights.fb.com/wp-content/uploads/2025/12/Israel-Palestine-HRDD-Implementation-update-2025.pdf>

³ Meta, Israel and Palestine Human Rights Due Diligence (2022), Recommendation 11: launched a Hebrew classifier that proactively detects and actions violating content in Hebrew: <https://about.fb.com/wp-content/uploads/2022/09/Meta-Response-Israel-and-Palestine-Due-Diligence-Exercise.pdf>

⁴ Meta, How Meta Is Preparing for Indian General Elections 2024: <https://about.fb.com/news/2024/03/how-meta-is-preparing-for-indian-general-elections-2024/>; Meta, How Meta is Preparing for Canada's 2025 Federal Election (20.3.2025): <https://about.fb.com/news/2025/03/how-meta-is-preparing-for-canadas-2025-federal-election/>;

⁵ How Meta Is Preparing for the 2026 US Midterm Elections <https://about.fb.com/news/2026/02/meta-prepares-for-2026-us-midterms>

⁶ Meta, MCA's WhatsApp Helpline: Curbing The Spread of AI-Generated Misinformation In India (19.2.2024): <https://about.fb.com/news/2024/02/mcas-whatsapp-helpline-curbing-the-spread-of-ai-generated-misinformation/>; Meta, How Meta Is Preparing for Indian General Elections 2024: <https://about.fb.com/news/2024/03/how-meta-is-preparing-for-indian-general-elections-2024/>

⁷ Protect the 2024 municipal elections in Brazil <https://about.fb.com/br/news/2024/04/nosso-trabalho-para-protoger-as-eleicoes-municipais-de-2024-no-brasil/>

⁸ Meta, Our Approach to Labeling AI-Generated Content and Manipulated Media (updated 2025): <https://about.fb.com/news/2024/04/metis-approach-to-labeling-ai-generated-content-and-manipulated-media/>

⁹ Meta, Germany 2025 Elections: <https://www.facebook.com/government-nonprofits/blog/germany-2025-elections>

¹⁰ Meta, New Tools to Support Independent Research – Meta Content Library & API (updated 26.9.2024): <https://about.fb.com/news/2023/11/new-tools-to-support-independent-research/>

¹¹ Meta, How Meta Is Preparing for Indian General Elections 2024 (WhatsApp forwarding limits): <https://about.fb.com/news/2024/03/how-meta-is-preparing-for-indian-general-elections-2024/>

¹² WhatsApp Business Policy – Political and Government Use: <https://business.whatsapp.com/policy>

מדיניות ומכשירים ייעודיים לבחירות שהופעלו מחוץ לישראל	פעילות ייעודית לזירה הישראלית
Messenger: מנגנון מוצר כללי להגבלת העברות – מאז עדכון 2025, המגבלה הכללית הוקלה, ומגבלה של העברה לעד חמש שיחות בכל פעם נשמרת רק לחשבונות מסוימים בעלי התנהגות חשודה או הונאתית, לצד סימון הודעות שהועברו פעמים רבות.¹³ Meta AI: בבחירות ארה"ב 2024 הופנו שאלות על אופן ההצבעה ותוצאות למקורות רשמיים.¹⁴	

פערי היערכות ומידע להיערכות המיוחדת עבור בחירות 2026 בישראל

- היערכות מקומית: מידע על צוות או מרכז בחירות ייעודי לישראל 2026, משאביו, שפות הכיסוי, זמני תגובה ומדדי ביצוע.
- Facebook / Instagram / Threads: פירוט מנגנוני האכיפה והסימון ב-Feed, Reels, Stories ותוכן ממומן של יוצרים; נתוני Ad Library ובדיקת שלמות הכיסוי בישראל.
- שותפים ודיווח: הרחבת Trusted Partners, בדיקת עובדות וערוץ הסלמה ייעודי עם מומחים וארגוני חברה אזרחית מקומיים.
- מחקר ושקיפות: אישור זכאות חוקרים בישראל ל-Meta Content Library/API ופירוט היקף הנתונים בכל שירות.
- WhatsApp: מענה נפרד על שיחות, קבוצות, Channels ו-Status; ערוץ בחירות ובדיקת עובדות מקומית; מדדי טיפול בהפצה המונית ונתונים מצרפיים תוך שמירה על פרטיות והצפנה.
- Messenger: פירוט מגבלות הפצה בתקופת הבחירות, טיפול בהתחזות ובהפצה מתואמת וערוץ הסלמה ישראלי.
- Meta AI: פירוט מנגנוני מענה לשאלות על הצבעה, מועמדים ותוצאות והפניה למקורות רשמיים בישראל.

¹³ Meta, Introducing a Forwarding Limit on Messenger (3.9.2020; updated 2.7.2025): <https://about.fb.com/news/2020/09/introducing-a-forwarding-limit-on-messenger/>

¹⁴ Meta, What We Saw on Our Platforms During 2024's Global Elections (Dec. 2024), including election-related Meta AI responses directing users to authoritative sources: <https://about.fb.com/news/2024/12/2024-global-elections-meta-platforms>

מיפוי מנגנוני הגנה על בחירות והשלמות נדרשות בשירותים דיגיטליים בישראל: Google YouTube, Search \ Ads, Gemini

המידע וההשלמות המבוקשים בישראל	מדיניות ומכשירים ייעודיים לבחירות מחוץ לישראל	פעילות ייעודית לזירה הישראלית
מידע על הקמת צוות או מרכז ייעודי לבחירות 2026 בישראל, משאביו והיקף אחריותו. הפעלת מנגנוני ההבלטה ותעדוף של מידע אמין בנושאי בחירות 2026 למשתמשים בישראל. התחייבות ומסירת פרטים על מנגנוני אכיפה ייעודיים להגנת הבחירות והתאמתם למגוון הלשוני בישראל, לרבות ביחס לשידורים חיים ו־YouTube Shorts. הרחבת נגישות נתונים למחקר ישראל: ב-Google Search ו-Gemini בהקשרי זיהוי ומניעת סיכונים לבחירות. נבקש מענה נפרד על תוצאות החיפוש, תשובות AI ומודעות: מקורות מידע רשמיים, ניטור מניפולציות, דיווח מועדף וגישת מחקר; וכן פירוט היקף הכיסוי של מאגר המודעות הפוליטיות בישראל (למשל מודעות בנושאים פוליטיים ללא מפלגה או מועמד).	<u>הבלטה ותעדוף מידע בחירות</u> ממקורות אמינים בתוצאות או בחלוניות מידע – Search, YouTube, Gemini (במדינות נבחרות).^{4, 5} <u>כלי הגנה ומשאבים טכנולוגיים</u>: לישויות בסיכון תקיפת סייבר או התחזות (ועדות בחירות, בעלי תפקידים, קמפיינרים, ועיתונאים).⁶ <u>שותפויות ייעודיות לבחירות עם ארגונים מקומיים</u>: תמיכה בקונסורציום של גופי חדשות ובודקי עובדות, מומחים) בכל השפות המרכזיות.⁷ <u>סימון תוכן סינתטי ב-Gemini וכלים לאימות תוכן ב-Gemini</u> (מדיניות גלובלית):⁸ SynthID, Content Credentials; הזיהוי ב-SynthID מוגבל לתוכן שנוצר בכלי Google. ב-Gemini יוצרים נדרשים לגלות שימוש ב-AI ליצירה או לשינוי של תוכן פוטוריאליסטי, והפלטפורמה מסמנת תוכן כזה, גם באמצעות זיהוי אוטומטי של תוכן שלא דווח. <u>אכיפה נגד מידע כוזב בבחירות ב-Gemini</u> (מדיניות גלובלית): איסור הטעיה לגבי מועד, מקום, אופן או זכאות להצבעה, טענות שקריות על כשירות מועמדים, ועידוד שיבוש הליכי הצבעה.⁹ <u>מחקר ושקיפות</u> (מדיניות גלובלית): גישת API למטא-דאטה של סרטונים.¹⁰	<u>פרסום פוליטי ממומן</u>: מדיניות ייעודית למודעות בחירות בישראל (מודעות המציגות מפלגה, נבחר ציבור או מועמד לכנסת, או שמפורסמות על ידם) הכוללת: מסלול אימות ייעודי למפרסמים וחובת גילוי של המפרסם בהתאם לדין הישראלי, הגבלות "טירגוט", חובת סימון של תוכן סינתטי או ערוך דיגיטלי, ופרסום המודעות במאגר השקיפות של Google למודעות פוליטיות בישראל.^{1, 2} <u>שותפים</u>: שדרוג שותפות הדיווח מול איגוד האינטרנט הישראלי בספטמבר 2025 (הצטרפות לתכנית Priority Flagger, לדיווח מזורז על שירותי Google מעבר ל-Gemini) בהיבטים השוטפים של מוגנות רשת,³ ללא הרחבה או תכניות דיווח ייעודיות בקשר לבחירות 2026.

¹ [Advertise Policies – Political Content](#): "Advertisers may run Israel Election Ads in Israel only if the advertiser completes Israel Election Ads verification."

² [Verification and Transparency – Election Verification Israel](#) ("During verification, advertisers must disclose whether they are a candidate, current elected officeholder, political party, or candidate list composed of one or more political parties for the Knesset. For these advertisers, all ads run in the verified account will be considered election ads. Advertisers must include disclaimers in their ads as required by Israeli law"). See also: [Ads Transparency – Israel](#)

³ [ISOC-IL Joins Google's Priority Flagger Program, Expanding Its Role in Protecting Israeli Users](#)

⁴ Google Public Policy (Elections), [supporting democratic processes around the world](#)

⁵ YouTube, [Election Information Panels](#)

⁶ Google Public Policy (Elections), [supporting democratic processes around the world](#)

⁷ Google, [Supporting the 2024 Indian General Election](#)

⁸ Gemini Apps, [Verify AI-generated images, videos, and audio](#)

⁹ [YouTube Policies – Misinformation – Elections Misinformation Policies](#):

¹⁰ YouTube Researcher Program, [How it works \(eligibility and participating countries\)](#)

TikTok: מיפוי מנגנוני הגנה על בחירות והשלמות נדרשות בשירותים דיגיטליים בישראל:

המידע וההשלמות המבוקשים בישראל	מדיניות ומכשירים ייעודיים לבחירות מחוץ לישראל	פעילות ייעודית לזירה הישראלית
• מידע על היקף הפעילות והמומחיות של מרכז הבחירות הייעודי והתאמתו לשפות המקומיות בישראל. שקיפות על פעילות האכיפה שנעשית בקשר לתהליך הבחירות לרבות פרסום נתונים ייעודיים לבחירות לפי סוג הפרה, ובחן הפרות טוהר הבחירות ומידע כוזב, תוכן AI מטעה. • פירוט אכיפת האיסור על פרסום פוליטי בישראל, לרבות תשלום ליוצרים מחוץ למערכת הפרסום, ופרסום נתוני זיהוי, הסרה וזמני טיפול. • חיזוק שיתופי פעולה עם חברה אזורית, ערוצי דיווח ייעודיים ומתועדפים לבחירות. לרבות קביעת מדדי ביצוע ייעודיים לבחירות, לרבות זמני תגובה, היקף זיהוי והסרה ושיעור זיהוי יזום. • שיתוף מידע ושקיפות בזירה הישראלית: הרחבת זכאות כלי המחקר לישראל ופרסום מדדי האכיפה ייעודיים שהוכרוזו לקראת בחירות 2026 בישראל.	<u>הפעלת מרכזי בחירות ייעודיים עם מומחיות מקומית:</u> בבחירות הפדרליות בקנדה ב-2025 הופעל מרכז בחירות באנגלית ובצרפתית, עם גישה למשאבי רשות הבחירות ב-16 שפות ילידיות; ההפניות אליו שולבו בתוכן רלוונטי. [3] <u>אכיפה ייעודית נגד איומי השפעה זרה ודיסאינפורמציה:</u> לקראת הבחירות לפרלמנט במולדובה ב-2025, החברה דיווחה על הסרת שלוש רשתות ובחן יותר מ-400 חשבונות שכוונו לקהלנים באירופה ומולדובה. [4] <u>סימון תוכן סינטי וכלים לאימות תוכן –</u> חובת סימון סינטי ואיסור על תוכן AI מטעה של אישי ציבור, לרבות הצגה כוזבת של תמיכה פוליטית. [5] תוכן לא מאומת עשוי לקבל תווית, אזהרה לפני שיתוף והחרגה מהמלצות. [6] <u>שותפים ודיווח:</u> שילוב אקדמיה, חברה אזרחית ויוצרים מקומיים בהסברת ההצבעה וביעוץ לשירות, וכן ערוצי דיווח ייעודיים לרשויות ולשותפים מקומיים. [3] [4] [7] <u>מחקר ושקיפות:</u> פרסום דוחות חודשיים על רשתות השפעה שהוסרו; Research Tools הזמינים לחוקרים בארה"ב, באירופה, בקנדה ובברזיל (ללא ישראל). [8]	בהודעה מיום 29.07.2026 הודיעה טיקטוק על הצעדים הבאים עבור תהליך בחירות 2026 בישראל: [1] הפעלת מרכז בחירות וצוות משימה ייעודי לבחירות הכולל מומחי סייבר, מומחים להתמודדות עם מידע כוזב וצוותי אמון ובטיחות. <u>פרסום ממומן:</u> יישום בישראל של מדיניות גלובלית על איסור פרסום פוליטי ממומן, לרבות באמצעות יוצרים. חשבונות של פוליטיקאים ומפלגות חסומים לכלי פרסום ומונטיזציה, וגיוס כספים אסור. חריג מוגבל מאפשר לגופים רשמיים האחראיים לבחירות לפרסם מידע רשמי על הצבעה והשתתפות בבחירות. [1] [2] <u>טיפול במידע שקרי והבלטת מידע אמין על סדרי הבחירות:</u> מדיניות הסרת מידע שקרי על מועדי ואופן הצבעה וחסימת חשבונות המתחזים לפוליטיקאים או לגופי תקשורת. לפי ההודעה, מרכז הבחירות מבוסס על מידע רשמי של ועדת הבחירות המרכזית ונגיש מתוך חיפושים ותכנים בנושא הבחירות. צוץ שיתוף פעולה עם הוועדה ובדיקת עובדות באמצעות רוייטרס. כאשר לא ניתן לאמת מידעית תוכן, תתווסף תווית "מידע שטרם אומת", לצד בקשה לשקול מחדש לפני שיתוף ומניעת המלצה בפיד ה-For You. כמו כן, מופעלים כלי C2PA וסימון מים. [1] ברבעון הראשון של 2026, הוסרו בישראל 551,341 סרטונים; כ-3% מהם הוסרו בגין הפרת כללי היושרה והאותנטיות, שכללו מידע שגוי ופגיעה ביושרת הבחירות. מבין ההפרות הללו, 99.1% הוסרו לפני שדווחו (לפי הודעת החברה על בסיס דוח אכיפת הנחיות הקהילה של החברה). <u>מחקר ושקיפות:</u> החברה הודיעה כי תפרסם נתוני אכיפה רבעוניים ביחס לתוכן בישראל, אך לא פרסמו מדדי ביצוע ייעודיים לבחירות לכל רכיבי ההיערכות כגון זמני תגובה או טיפול בדיווחים.

¹ טיקטוק מגנה על חווית קהילת המשתמשים במהלך הבחירות בישראל, 29.7.2026

² TikTok: [Politics, Governments, and Elections](#)

³ TikTok: [Protecting Election Integrity in Canada's 2025 Federal Election](#)

⁴ TikTok: [Protecting the integrity of TikTok during the parliamentary elections in](#); TikTok Launches In-app Center to protect platform during Philippine elections.

⁵ TikTok: [Protecting Election Integrity in Canada's 2025 Federal Election \(23.3.2025\)](#); [More ways to spot, shape and understand AI-generated content \(19.11.2025\)](#)

⁶ TikTok: [Protecting the integrity of TikTok during the parliamentary elections in Moldova \(21.8.2025\)](#)

⁷ TikTok: [TikTok Launches In-app Center to protect platform during elections and keep its Philippine community safe \(28.1.2025\)](#); [Our work to prepare for the 2024 European elections](#)

⁸ TikTok: [Protecting Election Integrity in Canada's 2025 Federal Election \(23.3.2025\)](#); [TikTok: Research Tools](#)

OpenAI / ChatGPT

Documented measures specific or directly relevant to Israel	Election policies and mechanisms documented globally or outside Israel	Information and additional measures requested for Israel
Dedicated election readiness: No public documentation of an Israel-specific 2026 election program was identified in the sources reviewed. OpenAI's 2026 framework documents US/Brazil measures and names Intelligence and Investigations, Safety, Security and Integrity teams for election abuse.¹ Israel-relevant threat enforcement: In 2024 OpenAI disrupted “Zero Zeno”: an Israel-origin operation generating English/Hebrew and India-election content. OpenAI reported disrupting some India-election activity less than 24 hours after it began.⁴ Advertising: OpenAI lists Israel for self-service Ads Manager access where the advertising and billing legal entity is based in Israel; this does not establish user-side ad delivery. Its policy prohibits political ads and excludes political-content conversations from ad placement.^{6, 7}	Election readiness and localized measures: OpenAI's 2024 election work combined safety, threat intelligence, legal, engineering and policy. In the US, ChatGPT routed voting queries to CanIVote.org and results to AP/Reuters/election authorities; for the EU Parliament, voting questions went to the official site.² In May 2024 OpenAI reported engagement with officials in Europe, India, South Korea and Mexico, plus civil society/academic roundtables.³ In 2026 OpenAI announced AP live vote counts for United States and Brazil. In the US, it announced a Democracy Works partnership for voting logistics, offered election-infrastructure cybersecurity support, and reported engagement with NASS/NASED. In Brazil, its TSE plan provides Kodex for authorities and a general reporting form.^{1, 9} Political campaigning and election misuse: OpenAI prohibits scaled campaign advocacy, lobbying, election interference and demobilization, including automated outreach and individualized message targeting; specified human-directed internal campaign uses remain permitted.⁵ Synthetic-content provenance: Supported OpenAI images use C2PA and SynthID; supported OpenAI audio uses SynthID. Its verifier detects signals that do not establish accuracy or context.⁸ Evaluation, research and transparency: OpenAI publishes political-bias evaluations, offers eligible researchers subsidized API access, and reports election-related abuse. These support transparency/research, not an election-specific researcher-data interface.^{1, 10, 11}	Local election readiness: Disclose Israel election-monitoring functions, timing, resources, Hebrew/Arabic expertise, escalation and Central Elections Committee coordination. Clarify whether US-style election cybersecurity support can be adapted for Israel. Authoritative election information: Implement an Israel-specific mechanism for voting procedures, eligibility, polling, candidate/list information and official results. Explain official-source/web-search use and correction of stale or fabricated answers; test in Hebrew/Arabic and, where practicable, Russian/Amharic. Political misuse: Explain enforcement of election-interference, demobilization, scaled messaging, automated outreach and targeting rules across ChatGPT, API and agents. Clarify Israeli ad delivery and political-ad-ban enforcement; specify provenance/verification coverage and limitations for political and synthetic content. Reporting and escalation: Provide a priority reporting route for the Central Elections Committee and qualified civil society organizations, with defined handling and urgent-response times. Research and transparency: Publish localized evaluation methods/results, enable independent testing of authoritative-source behavior, political objectivity, misuse safeguards and provenance, and publish feasible aggregate Israeli abuse/enforcement data.

Sources

- OpenAI, “Election information and safeguards in 2026” (27 May 2026): <https://openai.com/index/election-safeguards-2026/>
- OpenAI, “How OpenAI is approaching 2024 worldwide elections” (15 January 2024; updated 8 November 2024): <https://openai.com/index/how-openai-is-approaching-2024-worldwide-elections/>
- OpenAI, response to Senator Mark Warner on the Tech Accord to Combat Deceptive Use of AI in 2024 Elections (29 May 2024): https://www.warner.senate.gov/wp-content/uploads/media/doc/OpenAI_Response_MSC_Elections_Accord.pdf
- OpenAI, “Operation ‘Zero Zeno’: Israel-linked influence activity” (May 2024; accessed 29 September 2026): <https://openai.com/index/disrupting-malicious-uses-of-ai-zero-zeno/>
- OpenAI Help Center, “Political Campaigning Restrictions” (accessed 29 September 2026): <https://help.openai.com/en/articles/20001255-political-campaigning-restrictions>
- OpenAI, “Ad policies” (updated 10 September 2026): <https://openai.com/policies/ad-policies/>
- OpenAI Help Center, “Ads Manager Availability” (accessed 29 September 2026): <https://help.openai.com/en/articles/20001245-ads-manager-availability>
- OpenAI Help Center, “Provenance signals (Content Credentials, SynthID) in OpenAI-generated content” (accessed 29 September 2026): <https://help.openai.com/en/articles/8912793-provenance-signals-content-credentials-synthid-in-openai-generated-content>
- OpenAI Help Center, “OpenAI’s TSE Compliance Plan and Submitting Reports about Brazil’s 2026 Elections” (accessed 29 September 2026): <https://help.openai.com/en/articles/20001505-openais-tse-compliance-plan-and-submitting-reports-about-brazils-2026-elections>
- OpenAI, “Defining and evaluating political bias in LLMs” (9 October 2025): <https://openai.com/index/defining-and-evaluating-political-bias-in-llms/>
- OpenAI, “Researcher Access Program application” (accessed 29 September 2026): <https://openai.com/form/researcher-access-program/>

Anthropic / Claude

Documented measures specific or directly relevant to Israel	Election policies and mechanisms documented globally or outside Israel	Information and additional measures requested for Israel
No public documentation of Israel-specific measures or policies. Anthropic's recent update on safeguarding elections covers the US midterms and Brazil.¹ Global safeguards relevant to Israeli users: Usage Policy is enforced through Safeguards Team monitoring. It prohibits personalized voter targeting, deceptive movements, automated outreach at scale that conceals its artificial origin, deceptive/false election content, voter suppression through deception or intimidation, and disruption of voting/certification.³ Advertising: Anthropic states Claude will remain ad-free, without adjacent sponsored links, advertiser-influenced responses or unsolicited third-party placements.⁵	Election safeguards and testing: Anthropic uses classifiers, human review, Threat Intelligence, Policy Vulnerability Testing and election evaluations. April 2026 tests covered 600 prompts plus multi-turn/autonomous influence operations; results are company-reported.^{1, 8} Country and region-specific preparation: In 2024 Anthropic reported 12+ election-test rounds across India, South Africa, Mexico, UK, France, EU and US, plus authoritative-source routing in the US, EU, UK and France. Testing identified outdated/inaccurate answers and uneven responses; Anthropic then tightened election policies, updated system prompts, adding knowledge-limit warnings and routing users to official sources.² Authoritative information and currentness in 2026: For 2026 Anthropic announced US TurboVote/Democracy Works and later Brazil banners, and tests web-search triggering for time-sensitive election queries; US results are company-reported.¹ Political-use rules: In 2025 Anthropic replaced its blanket campaign/lobbying ban with rules targeting deceptive/disruptive democratic activity and voter/campaign targeting.^{3, 4} Synthetic-content provenance: Anthropic announced future Claude text watermarking globally at launch; supported files will carry C2PA credentials. Its detection API is in private preview, with broader access planned.⁶ Threat disruption and transparency: Anthropic's September 2026 report describes nine influence operations, several election-timed, and says findings improve detection. Its Transparency Hub publishes election-integrity safeguards/monitoring.^{7, 8}	Local election readiness: Identify Israel election-monitoring teams, Hebrew/Arabic expertise, timing, escalation and Central Elections Committee coordination. If no local team exists, explain how safeguards/threat Intelligence functions are adapted to Israel. Authoritative election information: Implement Israel election banner or official-source routing using Central Elections Committee information. Test search triggering, accuracy, cutoff handling and corrections in Hebrew/Arabic and, and when practicable, also Russian/Amharic. Political misuse: Explain enforcement against personalized voter targeting, deceptive campaigns, false election information, synthetic political content intended to deceive, and voter suppression through deception or intimidation, including when conducted through multi-turn or agentic workflows. Clarify watermarking/C2PA coverage and detection access for media, fact-checkers and researchers. Reporting, escalation and local partnerships: Provide a priority reporting route for the Central Elections Committee and qualified civil society organizations, with response times and local-expert testing. Research and transparency: Publish Israel-specific and/or multilingual evaluation methods and limits, distinguishing single- from multi-turn/agentic tests. Enable independent replication and feasible privacy-preserving Israel misuse/enforcement data.

Sources

- 1 Anthropic, "An update on our election safeguards" (24 April 2026): <https://www.anthropic.com/news/election-safeguards-update>
- 2 Anthropic, "Elections and AI in 2024: observations and learnings" (12 December 2024): <https://www.anthropic.com/news/elections-ai-2024>
- 3 Anthropic, "Usage Policy" (effective 15 September 2025; accessed 29 September 2026): <https://www.anthropic.com/legal/aup>
- 4 Anthropic, "Usage policy update" (15 August 2025; changes effective 15 September 2025): <https://www.anthropic.com/news/usage-policy-update>
- 5 Anthropic, "Claude is a space to think" (4 February 2026): <https://www.anthropic.com/news/claude-is-a-space-to-think>
- 6 Anthropic, "How Claude's text watermark works" (14 August 2026; updated 1 September 2026): <https://www.anthropic.com/news/claude-text-watermark>
- 7 Anthropic, "Detecting and countering misuse of AI: September 2026" (10 September 2026): <https://www.anthropic.com/threat-intelligence-report-september-2026>
- 8 Anthropic, "Anthropic's Transparency Hub," Election Integrity section (last updated 23 July 2026; accessed 29 September 2026): <https://www.anthropic.com/transparency/voluntary-commitments>

Mapping Election-Protection Mechanisms and Requested Measures for Digital Services in Israel:
X / Grok

Documented measures specific or directly relevant to Israel	Election policies and mechanisms documented globally or outside Israel	Information and additional measures requested for Israel
Political advertising: Israel is expressly covered. Political Content/Campaigning ads require certification by X; local-law compliance; for 90 days pre-election through election day, ads will disclose the publisher and contact details and meet Israel-specific rules; Political Ads Library applies.¹ Authenticity covers misleading synthetic/manipulated media (global policy).⁶ Election information integrity/authenticity: Global policy states that Civic Integrity covers misleading participation information, voter suppression/intimidation and reach-restricting escalation labels; Authenticity covers inauthentic accounts / behavior and misleading synthetic/manipulated/out-of-context media. No Israel-specific 2026 deployment publicly documented.² Israel-specific preparedness: No reviewed source documents 2026 Israel risk assessment, cross-functional election structure, X prompts, Grok official-source referral, or election reporting/research arrangements.	Election-specific operational preparedness: X rates each national election's risk (high/medium/low) to allocate resources, supported by a cross-functional election team. In 2026, it mapped 12+ risks with an electoral authority, prioritizing voting-period cases.³ Authoritative election information: X has used “Election Day prompts”, redirected Grok election queries to official sources after inaccurate answers (2024), and added homepage and search prompts to electoral-authority information (2026).⁴ Election-authority cooperation/reporting: X and the electoral authorities have run expedited referrals and a dedicated alert channel since 2024; a 2026 agreement added a named authority contact, training and rapid-response mechanisms, with separate tracks for judicial orders, authority alerts and user reports.⁵ Synthetic content & AI: X signed the voluntary 2024 AI Elections Accord (model risk, detection, provenance, transparency), but its commitments ran only through 2024. Enforcement & Community Notes: Community Notes add context but do not trigger enforcement; serious Civic Integrity cases may separately get reach-restricting labels. The Notes algorithm and data are public.⁷ Research & transparency: Under the EU Digital Services Act, vetted researchers can request X data on content, Notes, ads and moderation; X also publishes some recommendation code.⁸ Grok /X 2026 election plan, covers Grok, but its Grok safeguards are limited to general audiovisual generation and are not election-specific.⁹	Local preparedness: Disclose Israel's election risk rating and the cross-functional team assigned to it; map Israel-specific risks with the Central Elections Committee, prioritizing the voting period. Authority cooperation: Set up a dedicated Central Elections Committee contact, rapid-alert channel and expedited referrals; publish response times. Authoritative information: Deploy X prompts and Grok redirects to official Central Elections Committee information on voting, candidate lists and results, tested in Hebrew, Arabic, Russian and Amharic. Community Notes/enforcement: Report Notes coverage and speed in Hebrew and Arabic, and Israel-specific Civic Integrity and Authenticity enforcement, including reach-restricting labels. Research/transparency: Grant vetted Israeli researchers access to election data; publish a post-election report on enforcement and moderation during Israel's 2026 general elections. Adopt election-specific X\Grok safeguards against misinformation, impersonation and vote recommendations; promptly correct erroneous election answers.

¹ X, “[Political Content](#)” (current policy; Israel-specific rules and Political Ads Library).

² X, “[Civic Integrity](#)” (Nov. 2025); X, “[Authenticity](#)” (Apr. 2025).

³ X, [Government Affairs Team, UK Parliament written evidence](#) (22 Apr. 2024); [European Commission, German DSA election stress test](#) (31 Jan. 2025); TSE, [X 2026 election-compliance summary](#) (10 Sep. 2026).

⁴ X, [Government Affairs Team, UK Parliament written evidence](#) (22 Apr. 2024). [Minnesota Secretary of State, Grok election-query changes](#) (26 Aug. 2024); TSE, [X election-information cooperation agreement](#) (28 Jul. 2026).

⁵ TSE, [X 2024 Memorandum of Understanding](#); TSE, [X election cooperation agreement](#) (28 Jul. 2026); TSE, [X 2026 election-compliance summary](#) (10 Sep. 2026).

⁶ X, “[Political Content](#)”; AI Elections Accord, “[The Commitments](#)” (voluntary commitments through 2024; X signatory); X, “[Authenticity](#)” (Apr. 2025).

⁷ X, “[Civic Integrity](#)” (Nov. 2025); X, “[About Community Notes on X](#)”.

⁸ X, “[DSA Data Access for Vetted Researchers](#)”; X, “[About Community Notes on X](#)”; TSE, [X 2026 election-compliance summary](#) (10 Sep. 2026).

⁹ TSE, [X 2026 election-compliance summary](#) (10 Sep. 2026) the plan covers X and Grok and separately describes existing Grok audiovisual-generation safeguards.