

מסמך עזר לארגונים

טיפים לזיהוי דיפ-פייק

מה לבדוק בתמונה, בקול ובשיחה, ואילו נהלים מגינים על הארגון כשהעין לא מספיקה.

בתחילת 2024 השתתף עובד כספים בהונג קונג בשיחת וידאו עם מי שנראו ונשמעו כמו סמנכ"ל הכספים של החברה ועמיתים נוספים. לפי משטרת הונג קונג, כולם היו שחזורים מזויפים. העובד העביר כ-200 מיליון דולר הונג קונגי, כ-25.6 מיליון דולר אמריקאי, ב-15 העברות. החברה, Arup, אישרה ש"נעשה שימוש בקולות ובתמונות מזויפים".⁸

מה לבדוק בתמונה ובווידאו

בפרויקט Detect Fakes של MIT Media Lab מציעים לשים לב לשמונה אזורים:⁵

- **הפנים.** זיופים מתקדמים הם כמעט תמיד שינוי של הפנים.
 - **לחיים ומצח.** האם העור חלק מדי או מקומט מדי? האם גיל העור מתאים לגיל השיער והעיניים?
 - **עיניים וגבות.** האם הצללים מופיעים במקומות הצפויים?
 - **משקפיים.** האם יש השתקפות, האם היא חזקה מדי, והאם הזווית שלה משתנה כשהאדם זז?
 - **שיער פנים.** האם השפם, הפאות או הזקן נראים אמיתיים?
 - **שומות.** האם השומה נראית אמיתית?
 - **מצמוץ.** האם האדם ממצמץ מספיק, או יותר מדי?
 - **שפתיים.** האם תנועות השפתיים נראות טבעיות?
- ה-FBI מוסיף סימנים: ידיים או כפות רגליים מעוותות, שיניים או עיניים לא מציאותיות, פנים לא ברורות, משקפיים ותכשיטים לא מציאותיים, צללים לא מדויקים, סימני מים, השהיה, התאמת קול ותנועות לא מציאותיות.¹

מה לבדוק בקול ובשיחה

- ה-FBI ממליץ להקשיב לטון ולבחירת המילים, כדי להבחין בין שיחה אמיתית של אדם קרוב לבין קול ששוכפל באמצעות AI.¹
- לפי ה-FBI, נוכלים יוצרים קטעי קול קצרים בקולו של קרוב משפחה, במצב חירום לכאורה, עם בקשה לעזרה כספית מיידית או דרישת כופר.¹ מצב חירום ודחיפות הם סיבה לעצור ולאמת.
- בשיחות וידאו, ה-FBI תיעד מקרים שבהם תנועות השפתיים לא תאמו לגמרי לקול, ושיעול או עיטוש נשמעו בלי תנועה מתאימה בתמונה.²

מה לבדוק בריאיון עבודה מרחוק

- ב-2022 דיווח ה-FBI על עלייה בתלונות על שימוש בדיפ-פייק ובפרטים אישיים גנובים במועמדות למשרות מרחוק, בעיקר בתחומי IT ותוכנה. חלק מהמשרות כללו גישה למידע של לקוחות, למידע פיננסי ולמערכות הארגון.²
- באותם מקרים, בדיקות רקע לפני העסקה גילו שהפרטים שמסרו חלק מהמועמדים שייכים לאדם אחר.²
- ב-2025 דיווח ה-FBI שעובדי IT מצפון קוריאה השתמשו בבינה מלאכותית ובהחלפת פנים בריאיונות עבודה בוויידאו. אחת ההמלצות: לשאול שאלות "רכות" על פרטים ספציפיים של מיקום ורקע לימודי.³

הסייג החשוב

- במחקר שפורסם ב-iScience ב-2021, משתתפים לא הצליחו לזהות דיפ-פייק בצורה אמينة, והעלאת המודעות ותמריץ כספי לא שיפרו את הדיוק. המשתתפים נטו לסווג סרטונים מזויפים כאמיתיים, והעריכו יתר על המידה את היכולת שלהם לזהות אותם.⁶
 - ב-MIT מדגישים שאין סימן יחיד שמסגיר זיוף.⁵
 - ב-NIST מציינים פער בין דיוק גבוה של כלי זיהוי במחקר לבין הקושי להשתמש בהם בעולם האמיתי.⁷
- לכן סימן חשוב הוא סיבה לבדוק, אבל היעדר סימנים אינו מוכיח שהתוכן אמיתי. ההגנה העיקרית היא נהלים, כמפורט בעמוד הבא.**

נהלים שלא תלויים בעין

הוראה חריגה בטלפון או בווידאו

- ה-FBI ממליץ לנתק, לאתר בעצמכם את מספר הטלפון של הארגון שלכאורה פנה אליכם, ולהתקשר אליו ישירות.¹
- ה-FBI ממליץ למשפחות לקבוע מילה או ביטוי סודיים לאימות זהות.¹ אותו עיקרון ישים גם בארגון, בין מנהלים לבין מי שמבצע העברות כספים. זו התאמה של המסמך, ולא המלצה של ה-FBI לארגונים.
- העברת כסף או מידע רגיש מעל סף שנקבע מראש כדאי שתחייב אישור של אדם נוסף, בערוץ נפרד. במקרה של Arup, עובד אחד ביצע 15 העברות אחרי שיחת וידאו אחת.⁸

גיוס ועובדים מרחוק

- ה-FBI ממליץ לאמת זהות בריאיון, בקליטה ולאורך כל תקופת ההעסקה.³
- ה-FBI ממליץ לבצע כמה שיותר משלבי הגיוס והקליטה פנים אל פנים.³
- בדיקת רקע מצליבה את הפרטים שמסר המועמד עם מקורות עצמאיים. לפי ה-FBI, בדיקות כאלה גילו שהפרטים של חלק מהמועמדים שייכים לאדם אחר.²

בכירים ותקשורת רגישה

- לפי ה-NSA, ה-FBI וה-CISA, איומים של מדיה סינתטית כוללים התחזות למנהלים ולבעלי תפקידים בתחום הכספים. הם ממליצים לשקול אימות בזמן אמת, ולהגן על בכירים ועל התקשורת שלהם.⁴
- ה-FBI ממליץ לצמצם תמונות והקלטות קול ברשת, להגדיר חשבונות ברשתות החברתיות כפרטיים ולהגביל עוקבים לאנשים מוכרים. ככל שיש פחות חומר גלוי, קשה יותר לבנות זהות מזויפת.¹

היערכות ותגובה

- ה-NSA, ה-FBI וה-CISA ממליצים על שיתוף מידע, על תכנון ותרגול של תגובה לניסיונות ניצול, ועל הדרכת עובדים.⁴

לפני שמחליטים על סמך סרטון

סרטון שמתעד לכאורה עובד, מנהל או אירוע אינו מספיק לבדו. לפני תגובה כדאי לבדוק חמש שאלות:

- מקור המידע:** מי פרסם את הסרטון לראשונה, ומתי?
- זהות:** האם האדם בסרטון תואם לזהות הדיגיטלית שלו לאורך זמן?
- הצלבה:** האם יש ראיות בלתי תלויות לאירוע הנטען?
- דפוסי פעילות:** האם הופיעו לאחרונה חשבונות חדשים או שינויים בזהות?
- הקשר ומניעים:** מי עומד מאחורי ההפצה, והאם יש סימנים לניסיון להטעות?

גם מקורות גלויים יכולים להיות מזויפים, והיעדר ממצאים אינו מוכיח שאדם או סרטון הם אותנטיים.

המאמר המלא, עם הסרטון, במוקד הידע: profcheck.co/he/knowledge/deepfake-lineup-insider-threat

מקורות

1. ה-FBI (IC3), התרעה IC3-120324, 3 בדצמבר 2024. ic3.gov/PSA/2024/PSA241203
2. ה-FBI (IC3), התרעה IC3-062822-PSA, 28 ביוני 2022. ic3.gov/PSA/2022/PSA220628
3. ה-FBI (IC3), התרעה IC3-012325-PSA, 23 בינואר 2025. ic3.gov/PSA/2025/PSA250123
4. ה-NSA, ה-FBI וה-CISA, Contextualizing Deepfake Threats to Organizations, 12 בספטמבר 2023. nsa.gov/Press-Room/Press-Releases-Statements/Article/3523329
5. פרויקט Detect Fakes, MIT Media Lab, media.mit.edu/projects/detect-fakes/overview
6. מחקר: 2021, iScience 24(11), "Fooled twice: People cannot detect deepfakes but think they can", Köbis, Doležalová & Soraperra. doi.org/10.1016/j.isci.2021.103364
7. תוכנית Guardians of Forensic Evidence של ה-NIST. nist.gov/programs-projects/guardians-forensic-evidence
8. כתבה ב-CNN על Arup, 16 במאי 2024. edition.cnn.com/2024/05/16/tech/arup-deepfake-scam-loss-hong-kong-intl-hnk