القسم الرابع

التزييف والمعلومات المضللة

الحملات المضللة — التزييف العميق — الذكاء الاصطناعي والجريمة

الحملات المضللة الموجَّهة

الحملات المضللة أصبحت جزءاً من الحروب النفسية باستخدام التكنولوجيا — تستهدف الرأي العام والناشطات وصانعات السياسات على حدٍّ سواء.

أنواع المعلومات المضللة

- تلفيق المحتوى الكاذب تماماً
- التلاعب بمحتوى حقيقي — عناوين مضلِّلة أو صور مقتطعة من سياقها
- انتحال صفة — استخدام علامة تجارية لجهة موثوقة
- محتوى دقيق مقرون بسياق مغلوط
- السخرية تُعرَض على أنها حقيقة
- الروابط الزائفة التي لا تدعم ما تدّعيه

أدوات الحملات المضللة

- **الروبوتات:** برامج آلية قادرة على الوصول لـ 10,000 مستخدم وإيهامهم بمعارضة جماعية
- **الذكاء الاصطناعي:** يُنتج محتوىً مُقنعاً ومزيفاً من نصوصاً وصوراً
- **الحسابات الوهمية (اللجان الإلكترونية):** انتحال هوية أطراف أخرى للتلاعب بالرأي العام
- **الموقع الجغرافي:** استهداف دقيق لضحايا بعينهن
- **الإغراق الإعلامي:** رسائل متشابهة لإسكات الجهات المعارضة

الحملات المضللة الموجَّهة

كيف تحمين نفسك — طريقة SIFT للتحقق السريع

الحرف	الإجراء
S — Stop	توقّفي — لا تتفاعلي فوراً
I — Investigate	تحققي من المصدر
F — Find better sources	ابحثي عن مصادر أفضل
T — Trace claims	تتبّعي مصدر الادعاء الأصلي

تتعدد أدوات التحقق من الحملات المضللة الموجهة لتشمل تقنيات الكشف عن الحسابات الوهمية، تحليل الشبكات، والتحقق من الوسائط.

أدوات التحقق المجانية

- **فتبيّنوا (fatabyyano.com)** — للمحتوى العربي
- **مسبار (misbar.com)** — تحقق من الأخبار العربية
- **Google Reverse Image** — البحث العكسي عن الصور
- **Google Alerts** — مراقبة الشائعات عن منظمتك مجاناً

التحقق من المعلومات والصور

🛡️ قبل المشاركة — توقّفي 30 ثانية
السؤال قبل المشاركة ليس شكّاً — هو مسؤولية رقمية.

أسئلة تطرحينها على كل خبر تتلقينه

- من كتب هذا الخبر؟ هل المصدر موثوق؟
- متى نُشر؟ هل هو حديث أم قديم مُعاد تداوله؟
- لماذا يُشاركني الآن؟ ما الهدف؟
- هل هناك مصادر أخرى تؤكده؟
- هل يُثير مشاعر قوية جداً؟ (الإثارة المفرطة علامة تحذير)

التحقق من الصور

- **Google Reverse Image Search** — ابحثي عن مصدر الصورة الأصلية
- **TinEye.com** — تتبع أول ظهور للصورة على الإنترنت
- **ابحثي** عن تباعد الألوان أو العلامات المائية غير الاحترافية
- **تحققي من اسم النطاق** — قد يحتوي على أخطاء مقصودة مثل: gooogle.com

التحقق من الفيديوهات

- ابحثي عن عدم تزامن حركة الشفاه مع الصوت
- راقبي حدود الوجه — التشوهات عند الأذنين والأنف
- انتبهي لأنماط الإضاءة غير المتسقة بين إطار وآخر
- استخدمي أدوات كشف التزييف مثل: Hive Moderation أو Sensity.ai

التزييف العميق — Deepfakes

⚠ تهديد مباشر للنساء

التزييف العميق يُستخدَم بصورة متزايدة لإنشاء صور وفيديوهات مُزيَّفة للنساء في سياقات إباحية أو مُهينة دون موافقتهن. 98% من فيديوهات التزييف العميق الجنسية تستهدف النساء — Security Hero 2023. واكثر من 50% من النساء المصريات تعرضن للعنف الرقمي — مرصد الأزهر 2025.

ما هو التزييف العميق؟

بدأ التزييف العميق عام 2014 حين قدّم الباحث إيان غودفيلو فكرة الشبكات التوليدية التنافسية GANs. في 2017 انتشر على Reddit. منذ 2018 أصبح أداةً للاحتيال والابتزاز.

كيف يعمل التزييف العميق؟ — الشبكات التوليدية التنافسية (GANs)

تتكوّن من شبكتين عصبيتين: المولّد يصنع المحتوى المزيف، والمميّز يحدد إذا كان حقيقياً. يتنافسان حتى ينتج المولّد مقطعاً لا يستطيع المميّز اكتشافه.

أساليب التزييف

- **تبديل الوجوه:** استبدال وجه شخص بوجه آخر في الفيديو
- **تغيير الملامح:** لون البشرة، الشعر، تعبيرات الوجه
- **محاكاة حركة الشفاه:** تركيب صوت شخص على وجه آخر
- **التوليد الكامل بالذكاء الاصطناعي:** إنشاء شخصية كاملة غير حقيقية

التزييف العميق — Deepfakes

علامات تكشفين بها التزييف

- الرمش غير طبيعي أو غائب تماماً
- الشفاه لا تتزامن مع الصوت
- تغيرات في لون البشرة من إطار لآخر
- تشوهات عند الأذنين والأنف
- إضاءة غير متسقة بين الوجه والجسد
- الشخص يرفض الالتفات أو التحرك بشكل طبيعي
- جودة مختلفة بين منطقة الوجه وبقية الصورة

📖 شركة Arup — احتيال بالتزييف العميق 2024

في مطلع 2024، خسرت شركة Arup الهندسية 25 مليون دولار بعد مكالمة فيديو مزيّفة انتحل فيها المهاجمون صفة المدير المالي وموظفين آخرين. وفي نفس العام، استُخدم تزييف صوتي لانتحال صفة الرئيس بايدن في الانتخابات التمهيدية.

الحماية من التزييف العميق

كيف تحمين نفسك؟

- قبل مشاركة أي فيديو — خذي 30 ثانية للتحقق من المصدر
- لا تشاركي محتوى يستخدم صورتك حتى لإثبات أنه مُزيَّف
- أبلغي المنصة فوراً عن أي محتوى مزيَّف يستخدم وجهك أو صوتك
- سجِّلي في StopNCII.org إذا كان المحتوى جنسياً
- استشيري محامياً أو منظمة حقوقية متخصصة

📖 المستوى المجتمعي

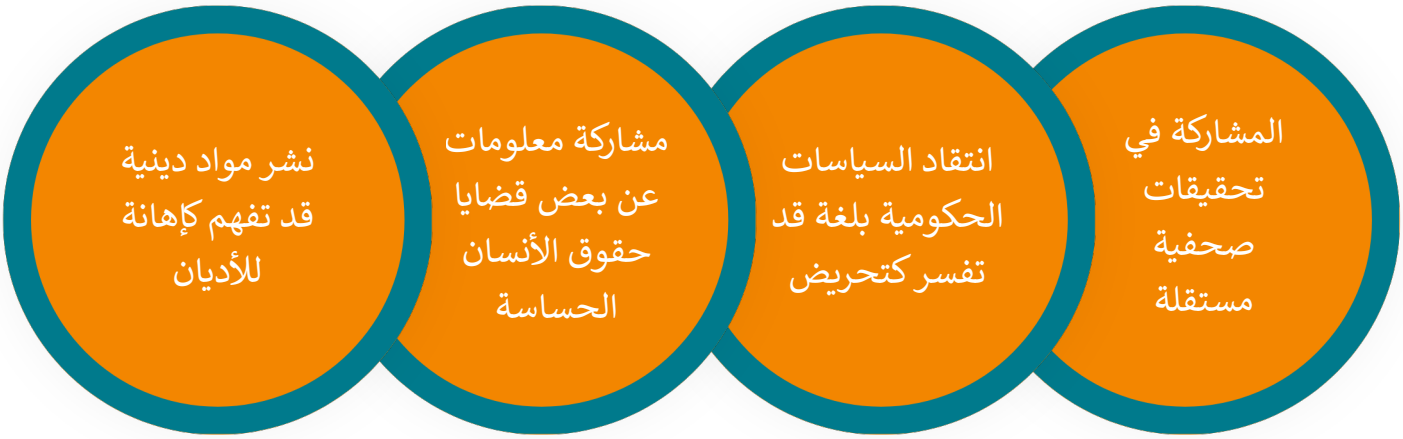
توصي منظمة WITNESS بخلق آليات والتزامات دولية مع الشركات للمساهمة في مساءلة إساءة استخدام الذكاء الاصطناعي وإنشاء محتوى فاضح دون موافقة، وتزويد الناشطات والصحفيات بأدوات الحماية اللازمة.

قضايا سلامة المحتوى والتعبير

متي يصبح التعبير الرقمي خطرا؟

التوازن الدقيق
حقك في التعبير عبر الانترنت مكفول قانونيا لكن بعض المحتوي يعرضك لمخاطر قانونية أو أمنية في السياق المصري. الوعي بهذا التوازن ضروري.

محتوي قد يعرضك لمخاطر



📖 **مورد مهم: دليل EFF لحرية التعبير الامنه**
تنشر منظمة EFF دليلا شاملا باللغة العربية حول كيفية ممارسة التعبير الرقمي بأمان

استراتيجيات التعبير الأمن

- ✓ استخدمي حسابات منفصلة: شخصي وعمل وناشط
- ✓ تعلمي استخدام ميزات "الاعدادات المخصصة" لكل منشور
- ✓ فكري مرتين في النشر المباشر للمعلومات الحساسة
- ✓ تواصلي مع مجموعات دعم موثوقة من نفس مجتمعك
- ✓ شاركي في تدريبات السلامة الرقمية الجماعية
- ✓ وثقي الانماط المتكررة وأبلغي المنتظمات الحقوقية
- ✓ تبادلي خبرات الحماية مع النساء الاخري

الأمان الجماعي أقوي من الفردي

بناء شبكات دعم مجتمعية تشترك في رصد التهديدات وتبادل المعلومات وتوثيق الأنماط يعزز الحماية بصورة جماعية.

الذكاء الاصطناعي والجريمة المنظَّمة

التحوّل الخطير

الذكاء الاصطناعي لم يخلق نوعًا جديدًا من الجرائم — بل جعل تنفيذها في متناول الجميع. رسالة تصيد كانت تستغرق 16 ساعة لكتابتها أصبحت تُنجَز في 5 دقائق.

أكثر من 90%
من إجمالي عمليات الاختراق الرقمي مدعومة بالذكاء الاصطناعي

88%
من هجمات التصيد تستهدف سرقة بيانات تسجيل الدخول - تحليل كاسبرسكي

أدى الاعتماد على الذكاء الاصطناعي إلى زيادة سرعة الهجمات الإلكترونية بمقدار 100 مرة

كيف يُعزِّز الذكاء الاصطناعي الهجمات؟

- يكتب رسائل تصيد مقنعة بلغة عربية سليمة خالية من الأخطاء
- يُنشئ فيديوهات وصوراً مزيّفة (التزييف العميق) لحملات الابتزاز
- يطوّر برمجيات خبيثة قادرة على تغيير سلوكها أثناء تنفيذ الهجوم
- ينتشر عبر تطبيقات مراسلة كواتساب لاستهداف المؤسسات

📖 حصان طروادة المصرفي — Trojan/Triada

فيروسات مدمجة في تطبيقات مزيفة أو أجهزة مباعة في السوق السوداء. تستهدف البيانات المصرفية وتعمل خِفية على الأندرويد. ارتفعت هجماتها على البنوك المصرية 186%، محتلةً المرتبة الثالثة في المنطقة. (Kaspersky 2026)

📖 برمجية Nexus — Android Banking Trojan

برمجية خبيثة من نوع حصان طروادة المصرفي تستهدف تطبيقات الهواتف الذكية التي تعمل بنظام الأندرويد لسرقة بيانات الدخول والمعلومات المالية. تنبيه: لا تخلطي بينها وبين عمليات حرب المعلومات الروسية-الأوكرانية — فهي أداة مختلفة تماماً.

الحماية من الجريمة الرقمية المُدعَّمة بالذكاء الاصطناعي

مؤشرات التهديد الناشئة 2026

- **التصيد بالذكاء الاصطناعي:** رسائل مخصَّصة تستخدم معلوماتك الشخصية الفعلية
- **تزييف المكالمات الصوتية:** انتحال صوت شخص تعرفينه لطلب المساعدة أو المال
- **هجمات NFC:** استهداف المحافظ الإلكترونية عبر تقنية الاتصال القريب
- **السوق السوداء:** هواتف أندرويد مباعة تحمل فيروسات مثبَّتة مسبقاً

⚠ هواتف الأندرويد المباعة مسبقاً

رُصد بيع هواتف أندرويد في السوق السوداء تحمل فيروسات مثبَّتة مسبقاً. اشتري دائماً من متاجر موثوقة وتحققي من خلوّ الجهاز من التطبيقات المشبوهة عند الشراء.

خطوات الحماية العملية

1. **استخدمي برنامج مكافحة فيروسات على هاتفك** — Bitdefender أو Kaspersky (نسخة مجانية)
2. **لا تثبتي تطبيقات خارج** Google Play أو App Store
3. **تحققي من الروابط** قبل النقر عبر virustotal.com
4. **لا تدفعي أموالاً بناءً على مكالمة صوتية** — حتى لو كانت لشخص تعرفينه
5. **لا تصدِّقي أي عرض وظيفي يبدو مثالياً جداً** — احتيال وظيفي شائع

للإبلاغ في مصر

- **مباحث الإنترنت** — للبلاغات الرسمية
  *يجب تقديم البلاغ في أسرع وقت، ويفضل ألا يتجاوز ثلاثة أشهر من تاريخ الواقع
- **Speak Up** — دعم شامل للعنف الرقمي
- **SMEX** — دعم السلامة الرقمية للناشطات والمدافعات
- **Access Now Helpline** — دعم السلامة الرقمية للناشطات والمدافعات