



مركز سميت للدراسات
SMT Studies Center

عام
الذكاء
الاصطناعي
AI
THE YEAR OF ARTIFICIAL INTELLIGENCE 2026



أزمة الحقيقة الثقة المؤسسية في زمن المحتوى الاصطناعي

مقدمة

لעقود من الزمن، بُنيت الثقة المؤسسية على افتراض أن "الرؤية تصديق". عملت وسائل الإعلام والحكومة والهيئات العلمية كـ"حراس للمعرفة" Epistemic Gatekeepers، تتحقق من المعلومات قبل وصولها إلى الجمهور. ومع ذلك، فقد دشنت عشرينيات القرن الحادي والعشرين عصر الحقيقة الاصطناعية "Synthetic Truth"، حيث أصبح التمييز بين المحتوى الذي ينتجه الإنسان والمحتوى الذي تنتجه الآلة غير مرئي عملياً.

بيان المشكلة

تُعرّف "أزمة الحقيقة" بضغطين متزامنين:

الانفجار من جانب العرض: يسمح الذكاء الاصطناعي بالإنتاج الضخم للمعلومات المضللة فائقة الواقعية التي تتجاوز سرعات التحقق من الحقائق التقليدية. التآكل من جانب الطلب: مع إدراك الجمهور أن أي شيء يمكن تزييفه، يبدأون في الشك بكل شيء، وهي ظاهرة تضر بالحقيقة أكثر من الأكاذيب نفسها.

التحول في "معادلة الحقيقة"

تاريخياً، كانت تكلفة الخداع عالي الجودة "مثل.. مقطع فيديو مزور أو حملة دعائية منسقة" عالية، وتتطلب موارد على مستوى الدولة. اليوم، تحولت "معادلة الحقيقة".. عندما يقترب حجم الضوضاء من اللانهاية، تغرق الإشارة "الحقيقة المؤسسية"، مما يؤدي إلى بيئة انعدام الثقة "Zero-Trust Environment".



رؤية من الوثائق: تحليل تقرير المخاطر العالمية للمنتدى الاقتصادي العالمي 2024 يحدد "المعلومات المضللة والتضليل المستمد من الذكاء الاصطناعي" كأكبر خطر عالمي خلال العامين المقبلين، ويرجع ذلك أساساً إلى أنه يقوض شرعية العمليات الديمقراطية "الانتخابات".

التآكل المؤسسي.. انهيار نموذج حارس البوابة

تاريخياً، عملت المؤسسات كمرشحات. لقد امتلكت "احتكار التحقق"، السلطة الاجتماعية والقانونية لإعلان ما هو "حقيقة". يكسر المحتوى الذي يولده الذكاء الاصطناعي "AIGC" هذا الاحتكار من خلال إنشاء حجم من الحقائق البديلة "Alternative Realities" التي لا تملك هذه المؤسسات الأدوات اللازمة لمعالجتها.

الصحافة.. من "المسودة الأولى للتاريخ" إلى "الضوضاء الخوارزمية"

• **مفارقة "السرعة مقابل الدقة":** يكشف تحليل الوثائق أن غرف الأخبار تتعرض لضغط شديد لاستخدام الذكاء الاصطناعي لمواكبة دورة المعلومات على مدار الساعة طوال أيام الأسبوع. ومع ذلك، يشير تقرير رويترز إلى عجز في الثقة "Trust Deficit" .. الجمهور يشك بشدة في الأخبار التي يولدها الذكاء الاصطناعي.

• **صعود مواقع الأخبار "الزائفة" "Pink Slime" NewsGuard:** يظهر تحليل تقارير NewsGuard زيادة بنسبة 1000% في مواقع "الأخبار" التي يولدها الذكاء الاصطناعي والتي تحاكي الصحافة المحلية. تفتقر هذه المواقع إلى المحررين البشريين ولكنها تستخدم الذكاء الاصطناعي عالي الدقة لإنتاج قصص متحيزة أو ملفقة.

المنهجية

لتحليل هذه الأزمة، تستخدم هذه الورقة تحليل الوثائق النوعي "QDA". سنقوم بتجميع النتائج من ثلاثة "مجموعات وثنائق" متميزة: وثنائق المخاطر والسياسات العالمية: مثل، تقرير المخاطر العالمية للمنتدى الاقتصادي العالمي 2024، قانون الاتحاد الأوروبي للذكاء الاصطناعي، مدونة قواعد السلوك للأمم المتحدة بشأن نزاهة المعلومات. هذه تكشف كيف يرى قادة العالم التهديد للاستقرار. استطلاعات وسائل الإعلام والرأي العام: مثل، تقرير معهد رويترز للأخبار الرقمية، مؤشر إيدلمان للثقة. هذه توفر بيانات تجريبية حول الاتجاه الحالي للثقة العامة. الأطر الأكاديمية والتقنية: مثل، عمل تشيسني وسيترون حول "مكافأة الكاذب"، ومعايير C2PA لأصل المحتوى. هذه تقدم الأدوات النظرية لفهم آليات الأزمة

آليات الأزمة.. "مكافأة الكاذب"

موضوع محوري في تحليل وثنائقنا هو مكافأة الكاذب "Liar's Dividend". يشير هذا المفهوم، الذي شاعه الباحثان القانونيان بوبي تشيسني ودانييل سيترتون، إلى أن الخطر الأكبر للذكاء الاصطناعي ليس التزييف العميق بحد ذاته، بل الشك الذي يوفره للأطراف المذبذبة.

• **الآلية:** في عالم توجد فيه التزييفات العميقة، يمكن لسياسي تم ضبطه في مقطع فيديو حقيقي ومضّر أن يدعي ببساطة: "هذا تزييف عميق تم إنشاؤه بواسطة الذكاء الاصطناعي".

• **التأثير المؤسسي:** هذا يخلق "سلاحاً ذا حدين". تفقد المؤسسات القدرة على مساءلة الأفراد لأن "دليل الواقع" أصبح الآن قابلاً للطعن بشكل دائم.

"التشطي المعرفي" للجمهور

النتيجة الأعمق في مؤشر إيدلمان للثقة (2024) هي أن "أزمة الحقيقة" تؤدي إلى التشطي المعرفي. **رؤية من الوثائق:** عندما لا تستطيع المؤسسات العمل كحراس بوابة، يتراجع الأفراد إلى "صوامع المعلومات" حيث يثقون فقط بالمصادر التي تتوافق مع تحيزاتهم المسبقة.

النتيجة: لم تعد الحقيقة بنية تحتية مشتركة؛ بل تصبح هوية قبلية. يسرع الذكاء الاصطناعي هذا من خلال تخصيص "الحقيقة" عبر خوارزميات التوصية، مما يضمن أن مواطنين اثنين في نفس المدينة قد يعيشان في "واقعين" مختلفين تماماً.

سلاح "التشطي المعرفي".. البعد الجيوسياسي لأزمة الحقيقة

لا تقتصر أزمة الحقيقة على التآكل الداخلي للمؤسسات، بل تمتد لتصبح أداة استراتيجية في صراع القوى العظمى، حيث انتقلنا من عصر "الحروب السيبرانية" التي تستهدف البنية التحتية التقنية "الخوادم والشبكات" إلى عصر الحروب المعرفية "Cognitive Warfare" التي تستهدف البنية التحتية الذهنية للمجتمعات.

• **سلاح "الإنهاء المعرفي":** في هذا السياق، يُستخدم المحتوى الاصطناعي من قبل الفاعلين الدوليين ليس فقط لنشر "رواية بديلة"، بل لخلق حالة من "العجز المعرفي" لدى شعوب الدول الخصمة. الهدف الجيوسياسي هنا هو الوصول بالجمهور إلى نقطة "الإنهاء" من الحقيقة، حيث يؤدي إغراق الفضاء الرقمي بتزييفات عميقة متقنة تستهدف رموز الدولة أو تشكك في نزاهة العمليات الديمقراطية "مثل الانتخابات" إلى تحطيم الثقة في "فكرة المؤسسة" ذاتها.

التأثير المؤسسي: تفقد الصحافة مكانتها كـ "حارس بوابة" لأنها لم تعد تستطيع ضمان مصداقية مصادرها. عندما يقوم "صحفي مواطن" بتحميل تزييف عميق لمنطقة حرب، وتقوم وسيلة إعلامية كبرى بنشره عن طريق الخطأ، تتضرر "علامة" المؤسسة التجارية بشكل دائم. لقد أصبحت "تكلفة الخطأ" وجودية.

القضاء.. نهاية مبدأ "الرؤية تصديق"

"دفاع التزييف العميق" في المحكمة: من النتائج الحاسمة في الأوراق القانونية ظهور "مكافأة الكاذب" كاستراتيجية إجرائية. يجادل محامو الدفاع بشكل متزايد بأن الأدلة الرقمية المدينة "الصوت/الفيديو" يمكن أن تكون مولدة بالذكاء الاصطناعي. التحليل: هذا يحول عبء الإثبات. فبدلاً من أن تتحدث الأدلة عن نفسها، يجب على المحكمة الآن إجراء "محاكمة داخل محاكمة" للتحقق من الأصالة التقنية لكل ملف رقمي.

هلوسة السوابق القضائية: يُظهر تحليل ملفات المحاكم الأميركية الأخيرة، مثل، قضية "Mata v. Avianca" محامين يستخدمون دون علمهم استشهادات قانونية "مهلوسة" مولدة بالذكاء الاصطناعي. هذا يقوض السلطة المعرفية للقانون؛ فإذا كان "منطق المحكمة" يعتمد على قضايا غير موجودة، فإن البنية القضائية بأكملها تفقد ثقة الجمهور.



اقتصاد السمعة.. الهوية هي الندرة الجديدة

في الاقتصاد الكلاسيكي، القيمة هي دالة للندرة. في عصر ما قبل الذكاء الاصطناعي، كانت "المعلومات" نادرة نسبياً وبالتالي كانت ذات قيمة. في عصر الذكاء الاصطناعي التوليدي، أصبحت المعلومات وفيرة للغاية و"عديمة التكلفة". ونتيجة لذلك، تنهار "قيمة المعلومات"، بينما ترتفع "قيمة المصدر" بشكل كبير.

"علاوة الأصالة".. الندرة في عصر الوفرة

يكشف تحليل مؤشر إيدلمان للثقة (2024) عن "فجوة ثقة" كبيرة بين الشركات التي تستخدم الذكاء الاصطناعي وتلك التي يُنظر إليها على أنها "بقيادة بشرية".

• **التحول:** ننتقل من نموذج ثقة "يركز على المحتوى" إلى نموذج ثقة "يركز على المصدر".

• **العلاوة:** مع إغراق المحتوى الذي يولده الذكاء الاصطناعي "AIGC" للسوق، يصبح المحتوى "من تأليف بشري" سلعة فاخرة. بدأت المؤسسات التي يمكنها ضمان عملية "إنسان في الحلقة" "HITL" في فرض "علاوة الأصالة".

• **التحليل:** تمامًا كما حولت علامة "العضوي" صناعة الأغذية، نشهد ظهور علامات "المعلومات العضوية". لم تعد الثقة موجودة في الرسالة، بل في نسب المرسل.

• **تقويض السيادة المعرفية:** تستخدم القوى العظمى الذكاء الاصطناعي التوليدي لشن عمليات تأثير "Influence Operations" منخفضة التكلفة وعالية النطاق، مما يسمح بالتدخل في الشؤون الداخلية للدول الناشئة دون إطلاق رصاصة واحدة. من خلال تزييف "أزمات وهمية" أو "تسريبات مؤسسية" مصطنعة، يتم دفع المجتمعات نحو الاستقطاب الحاد، مما يجعل الدولة المستهدفة مشلولة داخلياً وغير قادرة على اتخاذ قرارات سيادية بناءً على حقائق مشتركة.

• **النتيجة الاستراتيجية:** إن تحويل "الحقيقة" إلى ساحة معركة يعني أن الأمن القومي للدول بات مرتبطاً بقدرتها على حماية "الفضاء المعلوماتي" من التلوث الاصطناعي الخارجي. فالمؤسسات التي تفقد ثقة مواطنيها بسبب حملات التضليل العابرة للحدود تصبح "دولة هشة معرفياً"، مما يسهل اختراقها سياسياً واقتصادياً.

فجوة التحقق

يشير تحليل الوثائق إلى أننا نعيش حالياً في "فجوة تحقق". لقد تجاوزت تقنيتنا للإبداع "الذكاء الاصطناعي التوليدي" بكثير تقنيتنا للتحقق "العلامات المائية/C2PA". وحتى يتم سد هذه الفجوة، ستستمر المؤسسات في فقدان مكانتها كـ "حراس بوابة" أمام الحجم الهائل من الضوضاء الاصطناعية.

الهوية الرقمية كـ "الحدود السيادية" الجديدة

• **أزمة الهوية:** في عالم "التزييف العميق كخدمة"، تصبح القدرة على إثبات "أنا إنسان" هي الأصول الأكثر قيمة في الاقتصاد الرقمي.

• **صعود السيادة البيومترية:** يُظهر تحليل الوثائق اتجاهًا نحو "الهوية المرتبطة بالأجهزة". للحفاظ على الثقة، تتجه المؤسسات نحو نماذج الهوية السيادية "SSI" حيث ترتبط سمعة الشخص بمفتاح بيومتري أو تشفير لا يستطيع الذكاء الاصطناعي تكراره بسهولة.

• **خطر "الانقسام الهوياتي":** من النتائج الحاسمة في تحليل المنتدى الاقتصادي العالمي هو خطر ظهور نظام طبقي جديد.. أولئك الذين لديهم "هويات موثوقة" والذين يمكنهم الوصول إلى الخدمات المصرفية والتصويت ووسائل الإعلام الموثوقة مقابل "غير الموثقين"، الذين يتم ترحيلهم إلى طبقات الإنترنت "الموبوءة بالروبوتات".

المعادلة الجديدة للثقة

يمكننا تصور التحول في القيمة المؤسسية من خلال المنطق التالي:

• **ما قبل الذكاء الاصطناعي:** كانت "تكلفة الإنتاج" عالية "توظيف الصحفيين، أطقم التصوير، إلخ"، مما شكل حاجزاً طبيعياً أمام دخول المعلومات المضللة. التكلفة العالية أشارت إلى نية عالية.

• **ما بعد الذكاء الاصطناعي:** انخفضت "تكلفة الإنتاج" إلى ما يقرب من الصفر. لذلك، للحفاظ على نفس مستوى الثقة، يجب على المؤسسات زيادة شفافيتها "المصدر" وسلطتها "الهوية التي تم التحقق منها بشرياً" بشكل كبير.

من "الكشف" إلى "المصدر".. إطار عمل C2PA

• **فشل الكشف:** يشير تحليل الوثائق الفنية إلى أن "كاشفات الذكاء الاصطناعي" غير موثوقة إحصائياً وعرضة لـ "النتائج الإيجابية الخاطئة" معاقبة غير الناطقين باللغة الإنجليزية الأصليين أو الكتابة البشرية المنظمة.

• **محور المصدر:** يمثل معيار "C2PA"، المدعوم من "Microsoft" و "Adobe" و "NYT"، تحولاً أساسياً. بدلاً من محاولة "القبض على الكذب"، يركز على "إثبات الحقيقة" من خلال البيانات الوصفية المشفرة "Cryptographic Metadata".

• **التأثير المؤسسي:** بالنسبة لمؤسسة إخبارية أو هيئة حكومية، أصبحت "السمعة" مرتبطة الآن بتوقيعها الرقمي. إذا لم يكن للمستند "سلسلة عهدة" يمكن التحقق منها من الكاميرا إلى الشاشة، فإنه يُعامل بشكل متزايد على أنه "مذنب حتى تثبت براءته".



يشير تحليلنا إلى أن استراتيجية "الدفاع المتعمق" هي المسار الوحيد القابل للتطبيق للمضي قدماً.

الدرع التكنولوجي.. C2PA

"ملصقات تغذية المحتوى"

• آلية "بيانات اعتماد المحتوى": ينشئ معيار C2PA مساراً للبيانات الوصفية "مقاوماً للتلاعب". تماماً مثل "ملصق التغذية" على الطعام، فإنه يخبر المستخدم: من قام بإنشاء هذا؟ ما هي أدوات الذكاء الاصطناعي التي استخدمت؟ متى تم تحريره؟

• قيود "الثغرة التناظرية": يحدد تحليل الوثائق الفنية لأوراق عمل الأمن السيبراني "الثغرة التناظرية" كخلل حاسم. إذا التقط مستخدم صورة مادية لتزييف عميق يُعرض على شاشة، فإن السلسلة المشفرة تنقطع.

• التأثير المؤسسي: بينما يعد C2PA أداة قوية للمؤسسات ذات الثقة العالية "مثل نيويورك تايمز أو وكالة Associated Press"، إلا أنه لا يفعل الكثير لوقف المعلومات المضللة "من القاعدة إلى القمة" على المنصات المشفرة مثل "Whatsapp" أو "Telegram"، حيث غالباً ما يتم تجرييد البيانات الوصفية لأسباب تتعلق بالخصوصية أو الضغط.

الدرع الاجتماعي.. محو الأمية الإعلامية و"التفنيد المسبق"

• من التحقق من الحقائق إلى "التفنيد المسبق": التحقق التقليدي من الحقائق هو رد فعل، يأتي بعد أن تكون الكذبة قد انتشرت بالفعل. يشير تحليل تقارير منظمة التعاون الاقتصادي والتنمية إلى تحول نحو "التفنيد المسبق".

العلامة التجارية المؤسسية "البشرية فقط"

يشير تحليل تقارير "الاختبار الأحمر" لـ "OpenAI" و "إرشادات مقيم جودة البحث" من "Google" إلى أن المحتوى "الذي يركز على الإنسان" يتم إعطاؤه الأولوية بواسطة الخوارزميات لمكافحة "نظرية الإنترنت الميت".. فكرة أن الإنترنت سيصبح حلقة من الذكاء الاصطناعي يتحدث إلى الذكاء الاصطناعي.

• محور السمعة: بالنسبة لصناع القرار، يعني "اقتصاد السمعة" أن العلامة التجارية الشخصية والإرث المؤسسي هما الدفاعان الوحيدان ضد التقادم المدفوع بالذكاء الاصطناعي. "العلامة التجارية" هي في الأساس "وكيل ثقة"، اختصار للجمهور لتجنب العمل الشاق المتمثل في التحقق من كل قطعة محتوى يستهلكونها.

عصر "إثبات الشخصية"

تشير التحاليل إلى أننا ندخل عصر "إثبات الشخصية". المؤسسات التي تفشل في تبني معايير الهوية المشفرة والمصدر ستجد "سمعتها" تتآكل بسبب "الانجراف الاصطناعي"، التراكم البطيء للمحاكاة الساخرة والزيف والهوسات التي يولدها الذكاء الاصطناعي والتي ستطغى في النهاية على الصوت المؤسسي الأصلي.

الحلول التكنولوجية مقابل الحلول الاجتماعية.. بناء "المرونة المعرفية"

تُخاض المعركة ضد المعلومات المضللة التي يولدها الذكاء الاصطناعي على جبهتين.. البنية التحتية "كيف يتم تسليم المحتوى؟" والفكر "كيف يتم استهلاك المحتوى؟".

دراسة حالة.. "أزمة الحقيقة" في الممارسة العملية

لتجسيد المخاطر التي تمت مناقشتها، نستعرض حادثتين مفصليتين كشفتنا عن هشاشة النظام البيئي للمعلومات الحالي وقدرة المحتوى الاصطناعي على هز أركان الثقة المؤسسية في لحظات:

نموذج "الدفاع المتعمق"

بناءً على تجميعنا للوثائق، يمكننا نمذجة العلاقة بين الحلول التقنية والاجتماعية باستخدام معادلة المرونة:

الأصل التقني "الدرع": يقلل الجهد المطلوب للتحقق من الحقيقة.

• **المحو الأمية النقدية "الدرع الواقي":** يزيد من الشك المطلوب لرفض الكذب.

• **سرعة الخوارزمية "التهديد":** السرعة التي يتم بها تضخيم المحتوى الذي يولده الذكاء الاصطناعي بواسطة خوارزميات وسائل التواصل الاجتماعي.

خطر "السخرية الشاملة"

موضوع متكرر في وثائق معهد رويترز وإيدلمان هو خطر أن تأتي "المحو الأمية الإعلامية" بنتائج عكسية. إذا قيل للجمهور "كل شيء يمكن أن يكون تزييفاً عميقاً"، فقد يتوقفون عن تصديق أي شيء.

يحذر تحليل تقرير المخاطر العالمية للمنتدى الاقتصادي العالمي لعام 2024 من أن "السخرية الشاملة" هي في الواقع انتصار للجهات الفاعلة السيئة. إذا لم يتمكن الجمهور من التمييز بين مقطع فيديو حقيقي لجريمة حرب ومقطع فيديو تم إنشاؤه بواسطة الذكاء الاصطناعي، فسوف ينفصلون ببساطة عن الواقع.

الحل.. تشير الوثائق إلى أن الهدف لا ينبغي أن يكون "الشك"، بل الوكالة المعرفية "معرفة كيفية ومكان التحقق من المعلومات

الدرع الاجتماعي.. محو الأمية الإعلامية و"التفنيد المسبق"

• **من التحقق من الحقائق إلى "التفنيد المسبق":** التحقق التقليدي من الحقائق هو رد فعل، يأتي بعد أن تكون الكذبة قد انتشرت بالفعل. يشير تحليل تقارير منظمة التعاون الاقتصادي والتنمية إلى تحول نحو "التفنيد المسبق". من خلال تعليم الجمهور تقنيات خداع الذكاء الاصطناعي "على سبيل المثال، كيف تنشئ شبكة GAN وجهًا"، يصبحون أكثر مرونة لمحتوى الخداع.

• **استراتيجية "القراءة الجانبية":** يؤكد تحليل الوثائق الخاصة بالأطر التعليمية على "القراءة الجانبية"، مغادرة موقع مشبوه للتحقق من المصدر في مكان آخر، بدلاً من القراءة العمودية "تحليل محتوى الموقع نفسه".

• **فجوة محو الأمية:** التحدي الرئيسي الذي تم تحديده في تقرير اليونسكو العالمي لرصد التعليم لعام 2023 هو "فجوة محو الأمية". تتطور تقنية الذكاء الاصطناعي في غضون أسابيع، بينما تستغرق المناهج التعليمية الوطنية سنوات لتحديثها. وهذا يخلق "نافذة ضعف" للجمهور.



• **الأثر المؤسسي:** مثلت هذه الحادثة أول تطبيق واسع النطاق لتزييف الصوت "Voice Cloning" للتلاعب بإرادة الناخبين، مما وضع نزاهة العملية الانتخابية على المحك.

• **تحليل الاستجابة:** تحركت لجنة الاتصالات الفيدرالية "FCC" بسرعة لحظر المكالمات الآلية التي تستخدم أصواتاً مولدة بالذكاء الاصطناعي. تكمن خطورة هذه الحالة في أنها جسدت مفهوم "مكافأة الكاذب"؛ فإذا كان من السهل تزييف صوت الرئيس، فإن أي تصريح حقيقي مستقبلي يمكن إنكاره بسهولة بدعوى أنه "تزييف عميق".

تؤكد هاتان الحالتان أن المؤسسات، سواء كانت مالية أو سياسية، لم تعد تملك ترف الوقت للتحقق التقليدي. الاستنتاج الرئيسي هنا هو أن "السمعة المؤسسية" وحدها لم تعد كافية لصد الهجمات الاصطناعية، بل يجب دعمها بـ "بنية تحتية تقنية للتشفير والمصدر"، مثل معايير C2PA. لتمكين الجمهور من التمييز الفوري بين الخبر والضوضاء.

الخلاصة وتوصيات السياسة.. خارطة طريق للمرونة المعرفية

"أزمة الحقيقة" ليست مجرد خلل تكنولوجي؛ إنها فشل هيكل لنظامنا البيئي للمعلومات. يشير تحليلنا لتقارير المخاطر العالمية، والمعايير التقنية، والدراسات الإعلامية إلى أن "الثقة" لم يعد من الممكن افتراضها، بل يجب هندستها.

1. واقعة "انفجار البنتاغون".. اهتزاز الأسواق المالية

• **الحدث:** انتشرت صورة مولدة بالذكاء الاصطناعي تظهر دخاناً كثيفاً يتصاعد بجوار مبنى البنتاغون الأميركي. تمت مشاركة الصورة من قبل حسابات "موثقة" على منصة X، وانتقلت بسرعة البرق إلى وسائل إعلام دولية.

• **الأثر المؤسسي:** تسببت الصورة في انخفاض مؤقت ومفاجئ في مؤشر "S&P 500" في بورصة نيويورك، مما كشف عن قدرة "التزييف العميق" على إحداث ضرر اقتصادي حقيقي في دقائق.

• **تحليل الاستجابة:** اضطرت وزارة الدفاع الأميركية ومديرية الإطفاء في فرجينيا لإصدار بيانات نفي رسمية. هذه الواقعة أثبتت أن "سرعة الخوارزمية" تتفوق بمراحل على "سرعة التحقق المؤسسي"، وأن مجرد "النفي" لا يمحو الأثر النفسي والاقتصادي الفوري.

2. مكالمات "بايدن" الآلية في نيوهامبشير.. تهديد العملية الديمقراطية

• **الحدث:** تلقى آلاف الناخبين مكالمات هاتفية مسجلة بصوت يحاكي تماماً صوت الرئيس الأميركي جو بايدن، يطلب منهم عدم التصويت في الانتخابات التمهيديّة.

السياسة: يجب على المؤسسات العامة "المحاكم، الوكالات الحكومية، غرف الأخبار" اعتماد التوقيعات الرقمية التشفيرية لجميع الاتصالات الرسمية.

التنفيذ: يجب إنشاء "سلسلة حراسة قضائية" للأدلة الرقمية، حيث لا تكون أدلة الفيديو أو الصوت مقبولة إلا إذا كان يمكن التحقق من مصدرها وصولاً إلى جهاز التسجيل الأصلي.

الهدف: تحييد "مكاسب الكاذب" من خلال ضمان أن المحتوى "الحقيقي" يمكن التحقق منه بنفس سهولة اكتشاف المحتوى "المزيف".

3. الاحتكاك الخوارزمي

للمحتوى "غير المتحقق منه"

الوثيقة الأساسية: مدونة قواعد سلوك الأمم المتحدة بشأن نزاهة المعلومات.

السياسة: يجب إعادة هندسة خوارزميات وسائل التواصل الاجتماعي لإدخال "احتكاك التحقق". يجب تقييد المحتوى الذي يفتقر إلى بيانات المصدر الوصفية أو يأتي من حسابات "شبيهة بالروبوتات" غير موثقة من السرعة الفيروسية "قوائم الاتجاهات، التوصيات الخوارزمية".

التنفيذ: يجب على المنصات إعطاء الأولوية للمحتوى "الذي يشارك فيه البشر" في خلاصات الأخبار، مما يخلق بشكل فعال "مساراً سريعاً" للمعلومات الموثقة و"مساراً بطيئاً" للضوضاء غير الموثقة.

الهدف: رفع "تكلفة الخداع" من خلال جعل وصول المعلومات المضللة التي ينتجها الذكاء الاصطناعي بكميات كبيرة إلى جمهور عالمي أكثر صعوبة.

لاستعادة السلطة المؤسسية وثقة الجمهور، نقترح بنية تحتية للثقة متعددة الطبقات.

1. "تفويض المصدر"

الوثيقة الأساسية: قانون الذكاء الاصطناعي للاتحاد الأوروبي "التزامات الشفافية" والأمر التنفيذي 14110.

السياسة: يجب على الحكومات أن تفرض على جميع نماذج الذكاء الاصطناعي التوليدية عالية المخاطر "تلك القدرة على إنشاء بشر واقعيين أو استنساخ الأصوات" تضمين بيانات وصفية متوافقة مع C2PA وعلامات مائية غير مرئية تلقائياً.

التنفيذ: يجب مطالبة المنصات "وسائل التواصل الاجتماعي، محركات البحث" بعرض "بيانات اعتماد المحتوى" بتنسيق موحد وسهل الاستخدام، مثل شارة "إنسان موثق" أو "بمساعدة الذكاء الاصطناعي".

الهدف: القضاء على "إخفاء هوية المصدر" الذي يسمح للتزييف العميق بالانتشار دون مسار للمساءلة.

2. "السيادة الرقمية" المؤسسية

الوثيقة الأساسية: المنتدى الاقتصادي العالمي.. "مستقبل الهوية الرقمية" (2023).





"وليس التضليل نفسه". بدلاً من إخبار الناس بأن "هذا الخبر كاذب"، يتم شرح الآلية التي يُصنع بها الكذب "مثل: كيف يتم التلاعب بالعواطف، كيف تُستخدم التزييفات العميقة، أو كيف تعمل حسابات البوتات". عندما يفهم الجمهور "التكتيك"، يصبح لديهم حصانة تلقائية ضد أي محتوى مستقبلي يستخدم نفس الأسلوب.

2. آليات بناء المناعة المؤسسية الاستباقية

• حملات التوعية بالأنماط "Pattern Recognition": يجب على المؤسسات التعليمية والإعلامية إطلاق حملات تركز على "تفكيك الأسلوب". على سبيل المثال، شرح كيف يمكن لصورة مولدة بالذكاء الاصطناعي أن تفتقر إلى تفاصيل دقيقة "مثل أطراف الأصابع أو انعكاس الضوء"، مما يجعل المواطن "مدققاً أولياً" قبل أن تصل إليه المعلومة.

• الشفافية الراديكالية "Radical Transparency": تبني المؤسسات مناعة قوية عندما تسبق الشائعة بنشر الحقائق الخام. إذا كانت هناك أزمة متوقعة، فإن إعلان المؤسسة عن الاحتمالات والسيناريوهات بوضوح يخلق الثغرات التي قد يستغلها "المحتوى الاصطناعي" لاحقاً لملء الفراغ المعلوماتي.

• ألعاب المحاكاة والتفاعل: أثبتت الدراسات أن إشراك الجمهور في تجارب تفاعلية "مثل ألعاب رقمية تطلب من المستخدم صناعة خبر كاذب ليفهم كيف يتم التلاعب بالآخرين" ترفع من قدرته على كشف التضليل بنسبة تتجاوز 25%.

4. المبادرات الوطنية "للتفنيذ المسبق"

الوثيقة الأساسية: المبادئ التوجيهية لسياسة اليونسكو لمحو الأمية الإعلامية والمعلوماتية "MIL".
السياسة: تحويل أولويات التعليم الوطنية من "التحقق من الحقائق" إلى "التطعيم المعرفي".

التفنيذ: دمج "محو الأمية بالذكاء الاصطناعي" في التعليم من رياض الأطفال حتى الصف الثاني عشر وتعليم الكبار، مع التركيز على آليات الوسائط الاصطناعية. يتضمن ذلك تدريس تقنية "القراءة الجانبية" وعلم نفس "التحيز التأكيدي".

الهدف: بناء مواطنين يمتلكون "الوكالة المعرفية"، القدرة على التنقل في بيئة خالية من الثقة دون الوقوع في "السخرية الشاملة".

نحو استراتيجية المناعة الرقمية.. من "تصحيح الخبر" إلى "التمنيذ المعرفي"

تثبت التجارب أن تصحيح المعلومات "Fact-checking" بعد انتشارها غالباً ما يفشل في محو الأثر الأول "تأثير الحقيقة المستمرة". لذا، يجب على المؤسسات الانتقال إلى استراتيجية "التفنيذ المسبق"، وهي بناء "مناعة رقمية" لدى الجمهور قبل تعرضهم للمحتوى المضلل.

2. فلسفة التمنيذ الرقمي

تعتمد هذه الاستراتيجية على مبدأ اللقاح الطبي؛ حيث يتم تعريض الجمهور لـ "جرعات مخففة" من أساليب التضليل

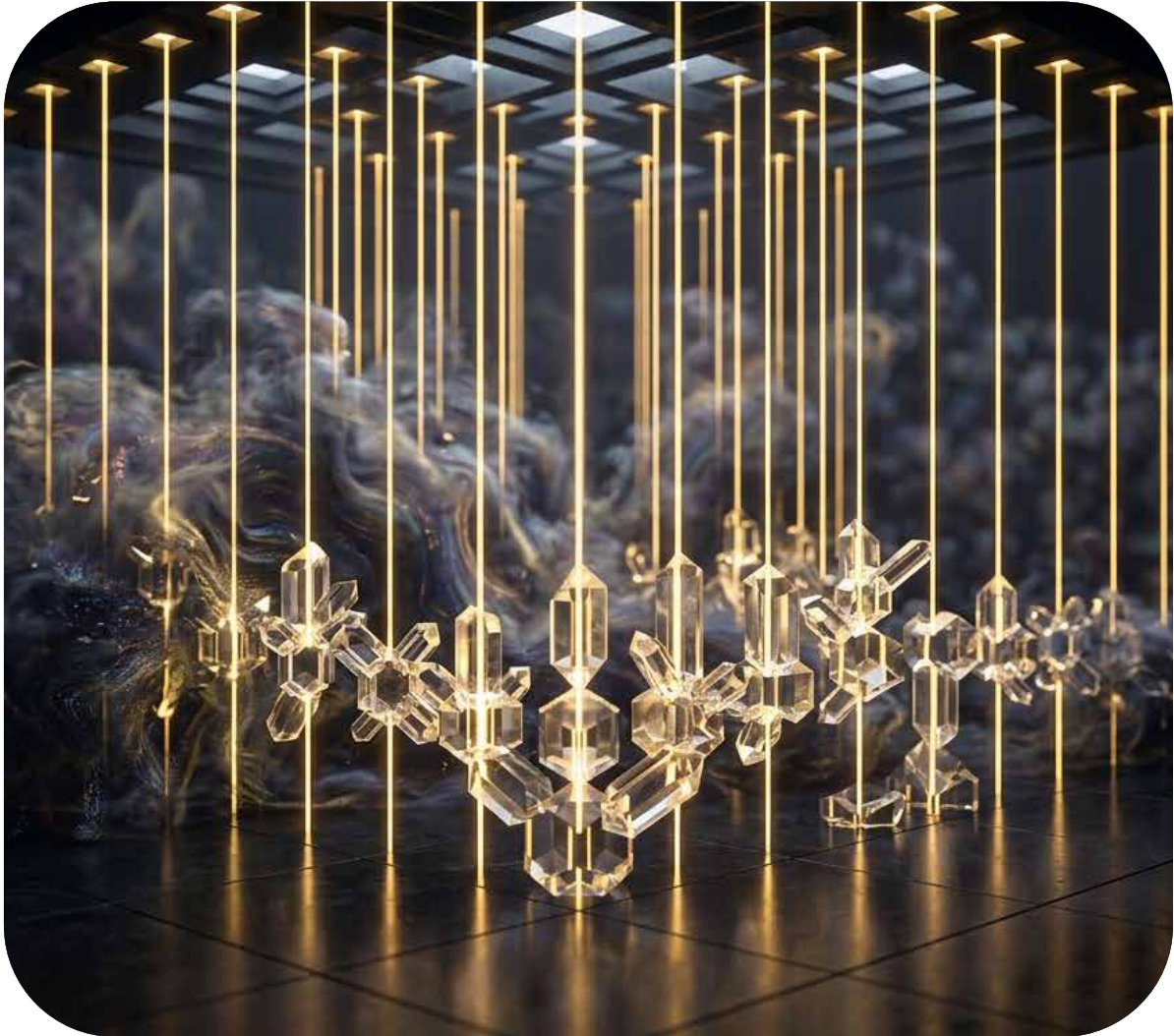
3. الانتقال من "النفي" إلى "التحصين"

إن هدف "التفنيد المسبق" هو تقليل الاعتماد على سرعة استجابة المؤسسة. ففي عالم يتحرك بسرعة الملي ثانية، لا يمكن للمؤسسة أن تلاحق كل كذبة، ولكن يمكنها بناء "عقل جمعي ناقد" يعمل كفلتر ذاتي. هذا التحول يحمي "الرأس مال الثقافي" للمجتمع ويجعل من محاولات التضليل الجيوسياسي عمليات مكلفة وغير مجدية.

مستقبل "معادلة الحقيقة"

تشير أبحاثنا إلى أن "أزمة الحقيقة" هي فترة انتقالية. نحن ننتقل من عصر الثقة الضمنية "حيث كنا نصدق ما نراه" إلى عصر التحقق الصريح "حيث نصدق ما يمكن إثباته".

إذا فشلت المؤسسات في تبني هذه الركائز، فستصبح عفا عليها الزمن بسبب "الانجراف الاصطناعي". ومع ذلك، إذا احتضنت اقتصاد السمعة والبنية التحتية للثقة، فيمكنها استعادة دورها كـ "حراس الواقع". إن بقاء الخطاب الديمقراطي لا يعتمد على "هزيمة الذكاء الاصطناعي"، بل على "التحقق من الإنسانية".



المراجع

- Adobe/C2PA. (2023). المواصفات الفنية لمصدر المحتوى وأصالتها.
<https://cutt.ly/lthq6ekw>
- تشيسني، ب.، وسيترون، د. (2019). التزييف العميق وفوضى المعلومات المضللة الجديدة: مكاسب الكاذب.
<https://cutt.ly/RtHq6Nwo>
- إيدلمان (2024). مقياس الثقة: الابتكار المكشوف.
<https://cutt.ly/ztHq6345>
- البرلمان الأوروبي (2024). قانون الذكاء الاصطناعي للاتحاد الأوروبي.
<https://cutt.ly/lthwqr9v>
- نيوزجارد (2024). صعود مواقع الأخبار والمعلومات التي يولدها الذكاء الاصطناعي.
<https://cutt.ly/KtHwql6y>
- معهد رويترز (2024). تقرير الأخبار الرقمية: الثقة في عصر الذكاء الاصطناعي.
<https://cutt.ly/MtHwqmPC>
- اليونسكو (2023). محو الأمية الإعلامية والمعلوماتية في عصر الذكاء الاصطناعي التوليدي.
<https://cutt.ly/TtHwqPTW>
- المنتدى الاقتصادي العالمي (2024). تقرير المخاطر العالمية: تهديد المعلومات المضللة.
<https://cutt.ly/WtHwqGgZ>



مركز سمث للدراسات
SMT Studies Center