الموجز السياساتي 8
بشأن خطتنا المشتركة

سلامة المعلومات على المنصات الرقمية

حزيران/يونيه 2023

الأمم المتحدة

مقدمة

توطئة

لن نستطيع التغلب على التحديات التي نواجهها إلا بتمتين التعاون الدولي. لذا، فإن مؤتمر القمة المعني بالمستقبل المزمع أن ينعقد في عام 2024 يتيح فرصة للاتفاق على حلول متعددة الأطراف من أجل غد أفضل، تعزز الحوكمة العالمية لصالح الأجيال الحاضرة والمقبلة على السواء (قرار الجمعية العامة 307/76). وقد دُعيتُ، بصفتي الأمين العام، إلى تقديم مساهمات في الأعمال التحضيرية لمؤتمر القمة في شكل توصيات عملية المنحى، استنادا إلى المقترحات الواردة في تقريري المعنون "خطتنا المشتركة" (A/75/982)، الذي كان إعداده في حد ذاته استجابة للإعلان المتعلق بالاحتفال بالذكرى السنوية الخامسة والسبعين لإنشاء الأمم المتحدة (قرار الجمعية العامة 1/75). وهذا الموجز السياساتي هو إحدى هذه المساهمات. وهو يتناول بإسهاب الأفكار التي اقترحت لأول مرة في خطتنا المشتركة، آخذا بالاعتبار التوجيهات التي قدمتها الدول الأعضاء في وقت لاحق وما جرى على مدى أكثر من عام من المشاورات الحكومية الدولية ومشاورات الجهات المتعددة صاحبة المصلحة، والتي تستمد جذورها من مقاصد ميثاق الأمم المتحدة ومبادئه والإعلان العالمي لحقوق الإنسان وصكوك دولية أخرى.

الغاية من هذا الموجز السياساتي

ويركز هذا الموجز السياساتي على الأخطار التي تتهدد سلامة المعلومات وأثرها على التقدم المحرز في القضايا العالمية والوطنية والمحلية. وكنت دعوت في خطتنا المشتركة إلى توافق مشترك تدعمه التجارب ويرتكز على الحقيقة والعلم والمعرفة. ولبلوغ هذه الغاية، يعرض هذا الموجز بإجمال المبادئ الممكن إدراجها ضمن مدونة لقواعد السلوك تساعد في توجيه ما تبذله الدول الأعضاء والمنصات الرقمية والأطراف الأخرى صاحبة المصلحة من جهود في سبيل جعل الفضاء الرقمي أكثر شمولا وأمانا للجميع، مع الدفاع المستميت عن الحق في حرية الرأي والتعبير والحق في الحصول على المعلومات. ويجري إعداد مدونة قواعد السلوك لتعزيز سلامة المعلومات على المنصات الرقمية ضمن سياق الأعمال التحضيرية لمؤتمر القمة المعني بالمستقبل. وأَمَلي أن توفر تلك المدونة المرجع الأسمى في توجيه العمل المضطلع به من أجل تعزيز سلامة المعلومات.

والمنصات الرقمية هي أدوات حاسمة غَيرتْ جذريا أوجه التفاعل الاجتماعي والثقافي والسياسي في كل مكان. فبواسطتها يستطيع الأشخاص الذين لهم وعي بالقضايا العالمية في جميع أنحاء المعمورة من التواصل بشأن المسائل التي تشد الاهتمام. وتساعد المنصات الأمم المتحدة على إبلاغ المعلومات إلى الناس والتواصل معهم مباشرة في سعيها إلى تحقيق السلام والكرامة والمساواة على كوكب سليم. وفتحت للناس أبواب الأمل في أوقات الأزمات والكفاح وأَسمَعتْ أصواتا لم تكن مسموعة من قبل ونفخت الحياة في الحركات العالمية.

ومع ذلك، كشفت هذه المنصات أنفسها أيضا عن جانب قاتم من المنظومة الرقمية. فقد مكنت من الانتشار السريع للأكاذيب والكراهية فتسببت في ضرر حقيقي على نطاق عالمي. وتضاءل التفاؤل بقدرة وسائل التواصل الاجتماعي على إفساح المجال أمام الناس للتواصل والتفاعل بسبب تفاقم المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وتحولها من ظاهرة هامشية في الفضاء الرقمي إلى ظاهرة فاشية. وهذه الظاهرة هي من الخطورة بحيث لا غلوَّ في التوكيد على جسامتها. فخطاب الكراهية والمعلومات المضللة اللذان يزدهران على وسائل التواصل الاجتماعي يمكن أن يقودا إلى العنف وإلى الموت[1]. وإن في القدرة على إشاعة المعلومات المضللة على نطاق واسع ونسف الحقائق الثابتة علميًا خطرًا وجوديًا يهدد البشرية (A/75/982، الفقرة 26) وهي أيضا خطرٌ على المؤسسات الديمقراطية وعلى حقوق الإنسان الأساسية. وقد تفاقمت هذه المخاطر بسبب التقدم السريع الذي حصل في مجال التكنولوجيا، كالذكاء الاصطناعي التوليدي. وتقوم الأمم المتحدة في جميع أنحاء العالم برصد خطر المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية على التقدم في تنفيذ أهداف التنمية المستدامة. ولقد أصبح جليا أن التعامل مع هذا الخطر بالطرق التقليدية لن يجدي نفعا.

قد يصعب في الممارسة العملية

التمييز بين

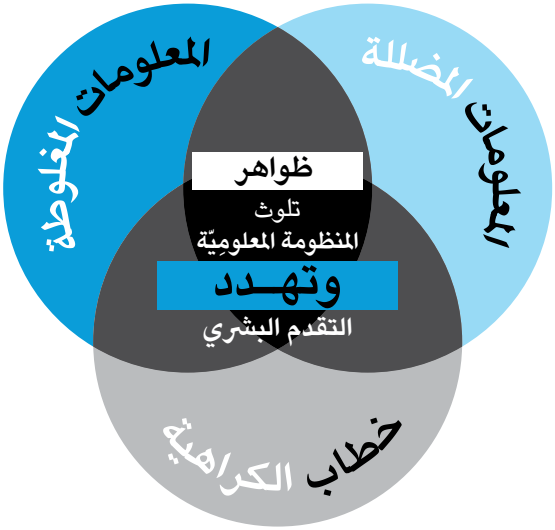
المعلومات المغلوطة

والمعلومات المضللة

ما معنى سلامة المعلومات؟

المقصود بسلامة المعلومات أن تكون المعلومات دقيقة ومتسقة وموثوقة. وتتعرض تلك السلامة للخطر بسبب المعلومات المضلِّلة والمعلومات المغلوطة وخطاب الكراهية. وعلى الرغم من عدم وجود تعاريف مقبولة عالميا لهذه المصطلحات، فقد وضعت كيانات الأمم المتحدة تعاريف عملية.

فالمقررة الخاصة المعنية بتعزيز وحماية الحق في حرية الرأي والتعبير تعرّف المعلومات المضللة بأنها "المعلومات الكاذبة التي تُنشر بنية إلحاق ضرر اجتماعي خطير"[2]. وتصف منظمة الأمم المتحدة للتربية والعلم والثقافة (اليونسكو) المعلومات المضللة بكونها المحتوى الخاطئ أو المضلل الذي يمكن أن يسبب ضررا محددا، بغض النظر عن الدوافع أو الوعي أو السلوكيات[3].



والعنصر الذي يأخذ به هذا الموجز السياساتي في الفرق بين المعلومات المغلوطة والمعلومات المضللة هو عنصر النية. فالمعلومات المضللة ليست معلومات غير دقيقة وحسب، وإنما هي أيضا معلومات تُشاع بنية الخداع وإيقاع الأذى بالغير. ويمكن أن تكون دولٌ أو أطراف أخرى من غير الدول مصدرا لإشاعة المعلومات المضللة في ظروف متعددة منها النزاعات المسلحة، ويمكن أن تؤثر تلك المعلومات في قاطبة ميادين التنمية، سواء تعلق الأمر بالسلام والأمن أم بحقوق الإنسان أم بالصحة العامة أم بالمساعدات الإنسانية أم بإجراءات مكافحة تغير المناخ.

أما المعلومات المغلوطة فهي المعلومات غير الدقيقة التي تشاع عن غير قصد ويتبادلها أناس مع غيرهم بنية حسنة دون أن يدركوا أنهم ينقلون الأكاذيب. ويمكن أن تكون المعلومات المغلوطة نابعة من المعلومات المضللة، حيث تُتخذ الأكاذيب والروايات المضللة سلاحا مع مرور الوقت، فتتسرب إلى الخطاب العام ويجري تناقلها عن غير وعي[4]. وقد يصعب في الممارسة العملية التمييز بين المعلومات المغلوطة والمعلومات المضللة.

وأما خطاب الكراهية فهو، حسب التعريف العملي الوارد في استراتيجية الأمم المتحدة وخطة عملها بشأن خطاب الكراهية، "أي نوع من أنواع التواصل في الكلام أو الكتابة أو السلوك الذي يهاجم شخصا أو جماعة أو يستعمل بحقهم لغة قدحية أو تمييزية على أساس هويتهم، أي بعبارة أخرى، على أساس انتمائهم الديني أو الإثني أو القومي أو العرقي أو أي عوامل أخرى تحدد الهوية من قبيل اللون أو الأصل أو نوع الجنس[5]".

والمعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وإن كانت تجمع بينها روابط إلا أنها ظواهر متمايزٌ بعضها عن بعض، مع وجود تداخلٍ فيما بينها في ميادين بعينها واختلافٍ في أساليب كشفها ومعالجتها والتخفيف من حدتها. ولكنها كلها تلوث المنظومة المعلومية وتهدد التقدم البشري[6].

على أن الأخطار التي تتهدد سلامة المعلومات ليست جديدة. إذ طالما أشيعت الأكاذيب والكراهية لتحقيق مكاسب سياسية أو مالية. ومع ذلك، فقد أصبح بالإمكان في العصر الرقمي القيام بذلك على نطاق لم يكن من الممكن تصوره من قبل. فأدوات الاتصال القوية أصبح في مقدورها حاليا أن تنشر أي محتوى في جميع أنحاء العالم في رمشة عين، وهذا يخلق مشكلة واسعة الانتشار لدرجة أن المنصات الإلكترونية نفسها تكون غير قادرة في بعض الأحيان على استيعاب مداها الكامل. ولا ينبغي أن يؤدي انتفاء تعاريف لهذه المصطلحات متفق عليها على الصعيد الحكومي إلى أن نظل مكتوفي الأيدي. فيجب أن نفعل كل ما في وسعنا للحد من الأضرار التي تسببها.

سلامة المعلومات والمنصات الرقمية

ينبغي أن تكون المنصات الرقمية طرفا أساسيا في السعي إلى تعزيز سلامة المعلومات. وعلى الرغم من أن بعض وسائل الإعلام التقليدية يمكن أن يكون أيضا منبعا للمعلومات المغلوطة والمعلومات المضللة، فإن سرعة انتشار تلك المعلومات وحجمها وذيوعها عبر القنوات الرقمية تستدعي اتخاذ تدابير عاجلة ومخصوصة لمواجهتها. ويشير مصطلح "المنصة الرقمية" في سياق هذا الموجز إلى خدمة رقمية تسهل التفاعل بين اثنين أو أكثر من المستخدمين، وتغطي مجموعة واسعة من الأنشطة، من وسائل التواصل الاجتماعي ومحركات البحث إلى التطبيقات الخاصة بالمراسلة. وهي تقوم عادة بجمع بيانات تتعلق بمستخدميها وتفاعلاتهم[7].

وتُختلق المعلومات المغلوطة والمعلومات المضللة من قبل مجموعة واسعة من الأطراف بدوافع مختلفة ويكون في مقدور تلك الأطراف عموما أن تبقي هويتها محجوبة خلف الأستار. وقد استَغلت حملات التضليل المنسقة التي قادتها دول وأطراف أخرى من غير الدول مواطن الضعف في النظم الرقمية لإشاعة الروايات الضارة، وكان لذلك تداعيات خطيرة.

واتخذ العديد من الدول مبادرات لوضع ضوابط تنظيمية للمنصات الرقمية، حيث جرى اعتماد أو النظر في اعتماد ما لا يقل عن 70 قانونا في السنوات الأربع الماضية[8]. وعادة ما تنطوي النهج التشريعية في جوهرها على نطاق ضيق من سبل الانتصاف لتحديد المحتوى الضار وإزالته. وبالتركيز على إزالة المحتوى الضار، وضع بعض الدول تشريعات معيبة وفضفاضة أدت في الواقع إلى إخراس "الكلام المشمول بالحماية"، الذي يأذن به القانون الدولي. وقد تفتقر التدابير الأخرى المتخذة، مثل التعطيل الشامل لإمكانية الاتصال بالإنترنت وحظر المنصات، إلى الأساس القانوني وتنتهك حقوق الإنسان.

واتخذ العديد من الدول والشخصيات السياسية من مخاوف مزعومة بشأن سلامة المعلومات ذريعة لتقييد الوصول إلى المعلومات، والقدح في التقارير وتقييدها، واستهداف الصحفيين والمعارضين[9]. وقامت جهات تابعة لدول بالضغط أيضا على المنصات لتنفيذ طلباتها تحت ستار التصدي للمعلومات المغلوطة والمعلومات المضللة[10]. وشدد خبراء حرية التعبير على أنه يقع على عاتق الجهات الفاعلة الحكومية واجب خاص في هذا السياق وأنه "لا ينبغي لها أن تقدم معلومات مغلوطة أو أن ترعاها أو تشجعها أو تزيد في نشرها" (A/77/287، الفقرة 45).

وتتطلب المخاطر الكامنة في تنظيم التعبير نهجا تُبذل العناية في تصميمه ويكون مستوفيا لشروط الشرعية والضرورة والتناسب بموجب قانون حقوق الإنسان، حتى عندما يكون هناك غرض مشروع تقتضيه المصلحة العامة (المرجع نفسه، الفقرة 42).

والمعلومات المضللة هي أيضا من الممارسات التي تلجأ إليها كبريات الشركات التجارية. فمن المصادر الرئيسية لنشر المحتوى المزيف والمضلل شركات العلاقات العامة التي يكون دورها هو تشويه سمعة الخصوم وكذلك تلك التي يكون لها في خدمات العلاقات العامة حضور نافذ والتي تتعاقد معها الدول والشخصيات السياسية والقطاع الخاص[11]. ومن بين أساليب المناورة التي يتم اللجوء إليها نشرُ محتوى في نسخ مزيفة من مواقع إخبارية لجعل المقالات تبدو وكأنها من مصادر شرعية[12]. فهذه الممارسات الغامضة يكون من الصعب جدا تعقبها والبحث عنها بحيث إن الحجم الحقيقي للمشكلة غير معروف. على أن الأفراد هم أيضا ينشرون ادعاءات كاذبة لبيع المنتجات أو الخدمات من أجل الربح، وغالبا ما يستهدفون الفئات الضعيفة في أوقات الأزمات أو انعدام الأمن.

ويتوقف النهج السائد في نماذج الأعمال الحالية التي تنتهجها معظم المنصات الرقمية على "الاقتصاد القائم على إثارة اهتمام المستخدمين". إذ يجري تصميم الخوارزميات لإعطاء الأولوية للمحتوى الذي يشد انتباه المستخدمين، ومن ثم زيادة المشاركة والإيرادات المتأتية من الإعلانات إلى أقصى حد ممكن. وكثيرا ما يكون المحتوى غير الدقيق والبغيض المصمم لاستقطاب المستخدمين وتهييج المشاعر هو الذي يولد أكبر قدر من المشاركة، وقد صار من المعروف نتيجة ذلك أن الخوارزميات تصب في مصلحة المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وتُضخمها[13].

وأمام انخفاض عائدات الإعلانات، تبحث المنصات الرقمية عن طرق بديلة لتحقيق الربح تتجاوز الاقتصاد القائم على إثارة اهتمام المستخدمين. وعلى سبيل المثال، فإن الخطط التي وضعتها المنصات الرقمية للبصم على صدقية الحسابات مقابل رسوم تُدفع، حيث يمكن لأصحاب تلك الحسابات شراء ختم إقرار سابق الاستعمال للدلالة على صدقية هويتهم، قد أثارت مخاوف فعلية فيما يتعلق بسلامة المعلومات بالنظر إلى احتمال إساءة استخدامها من قبل مُرَوّجي المعلومات المضللة[14].

بلايين

الناس يستخدمون
وسائل التواصل الاجتماعي

المصدر: KepiosDigital 2023 Global Overview Report

فيسبوك
3,0 بلايين

إنستاغرام
2,0 بلايين

واتساب
2,0 بلايين

يوتيوب
2,5 بلايين

تويتر
0,6 بليون

فيسبوك ميسينجر
0,9 بليون

سناب شات
0,6 بليون

دُوين
0,7 بليون

كوايشو
0,6 بليون

تيك توك
1,1 بليون

بينترست
0,4 بليون

سينو ويبو
0,6 بليون

وي تشات
1,3 بليون

كيو كيو
0,6 بليون

تيليغرام
0,7 بليون

ما هو الإطار القانوني الدولي ذو الصلة بالموضوع؟

يجب أن يرتكز تعزيز سلامة المعلومات ارتكازا تاما على القواعد والمعايير الدولية التي لها صلة بالموضوع، بما في ذلك قانون حقوق الإنسان ومبدأ السيادة وعدم التدخل في الشؤون الداخلية. وفي آب/أغسطس 2022، أحلْتُ إلى الجمعية العامة تقريرا بعنوان "مكافحة التضليل الإعلامي من أجل تعزيز وحماية حقوق الإنسان والحريات الأساسية"[15]. وبسطت في ذلك التقرير القانون الدولي لحقوق الإنسان المنطبق على المعلومات المضللة، بما في ذلك الإعلان العالمي لحقوق الإنسان والعهد الدولي الخاص بالحقوق المدنية والسياسية. ومن مقتضيات ذينك الصكين الدوليين أن حرية التعبير حق لكل شخص[16].

فالمادة 19 من الإعلان العالمي لحقوق الإنسان والمادة 19 (2) من العهد الدولي الخاص بالحقوق المدنية والسياسية تحميان الحق في حرية التعبير، وذلك يشمل حرية التماس جميع أنواع المعلومات والأفكار وتلقيها ونقلها، بصرف النظر عن الحدود، وبأية وسيلة كانت. ولا يقتصر حق الإنسان في حرية التعبير على المعلومات التي تُتلَقى بشكل إيجابي(A/77/287، الفقرة 13). أما حرية الإعلام، المرتبطة بحرية التعبير، فهي في حد ذاتها حق من الحقوق. وقد ذكرت الجمعية العامة ما يلي: "إن حرية الإعلام حق أساسي من حقوق الإنسان وهي معيار جميع الحريات التي تكرَّس لها الأمم المتحدة" (المرجع نفسه، الفقرة 14). ويمكن إخضاع حرية التعبير والحصول على المعلومات لقيود محددة تستوفي معايير خاصة مبسوطة في المادة 19 (3) من العهد[17]. ولا يمكن للدول أن تضيف أسسا أخرى أو تقيد حرية التعبير بما يتجاوز القيود المسموح بها بموجب القانون الدولي.

وتوفر خطة عمل الرباط بشأن حظر الدعوة إلى الكراهية القومية أو العرقية أو الدينية التي تشكل تحريضا على التمييز أو العداوة أو العنف، المعتمدة في عام 2012، إرشادات قانونية وسياساتية عملية للدول بشأن أفضل السبل لتنفيذ المادة 20 (2) من العهد والمادة 4 من الاتفاقية الدولية للقضاء على جميع أشكال التمييز العنصري، التي تحظر أشكالا معينة من خطاب الكراهية. وقد استعملت الدول الأعضاء بالفعل خطة عمل الرباط في سياقات مختلفة[18].

وكان خطاب الكراهية إرهاصا لجرائم فظيعة منها الإبادة الجماعية. وتحظر اتفاقية منع جريمة الإبادة الجماعية والمعاقبة عليها لعام 1948 "التحريض المباشر والعلني على ارتكاب الإبادة الجماعية".

وشددت الجمعية العامة في القرار 227/76، الذي اتخذته في عام 2021، على أن جميع أشكال التضليل الإعلامي يمكن أن تؤثر سلبا على التمتع بحقوق الإنسان والحريات الأساسية، وكذلك على تحقيق أهداف التنمية المستدامة. وكذلك أكد مجلس حقوق الإنسان في قراره 21/49، المعتمد في عام 2022، أن التضليل الإعلامي يمكن أن يؤثر سلباً على التمتع بجميع حقوق الإنسان وإعمالها.

كان خطاب الكراهية إرهاصا لجرائم فظيعة منها الإبادة الجماعية

قال 75 في المائة

من حفظة السلام التابعين للأمم المتحدة

إن المعلومات المغلوطة والمعلومات المضللة

تؤثر على

سلامتهم وأمنهم

استقصاء الأمم المتحدة الداخلي لعام 2022

ما الضرر الناجم عن إشاعة المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية على الإنترنت؟

تثير إشاعة المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية على الإنترنت قلقا حقيقيا لدى الجمهور العالمي. فقد وجدت دراسة لبيانات مستقاة من أجوبة مشاركين في استقصاءات في 142 دولة أن 58,5 في المائة من الأشخاص الذين يرتادون الإنترنت ووسائل التواصل الاجتماعي بانتظام في جميع أنحاء العالم قلقون من خطر التعرض للمعلومات المغلوطة على الإنترنت، وأشار شباب وأناس من الشرائح المنخفضة الدخل أنهم يعترون أنفسهم أكثر عرضة لذلك بدرجة كبيرة[19]. ولأن شباب اليوم قد ولدوا في العصر الرقمي فهم على الأرجح أكثر ارتيادا للإنترنت من بقية السكان، ومن ثم فهم الجيل الأكثر اتصالا بالإنترنت في التاريخ[20]. ويقدر أن نسبة الأطفال الذين يتصلون بالإنترنت لأول مرة في حياتهم في جميع أنحاء العالم هي طفلٌ واحد في كل نصف ثانية، وهذا ما يجعلهم عرضة لخطر خطاب الكراهية والأذى على الإنترنت، مما يؤثر في بعض الأحيان على صحتهم العقلية[21].

ويمكن معاينة آثار المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية على الإنترنت في جميع أنحاء العالم، بما في ذلك في مجالات الصحة وإجراءات مكافحة تغير المناخ والديمقراطية والانتخابات والمساواة بين الجنسين والأمن والاستجابة الإنسانية. واعتبر 75 في المائة من المكاتب القطرية التابعة لبرنامج الأمم المتحدة الإنمائي في استقصاء أجري في عام 2021 أن المعلومات الزائفة هي مصدر قلق كبير. ولها آثار خطيرة على الثقة والأمان والديمقراطية والتنمية المستدامة، كما تبين من استعراض أجري مؤخرا بتكليف من اليونسكو وشمل أكثر من 800 وثيقة من الوثائق الأكاديمية والصحفية وتلك المتعلقة بالمجتمع المدني والشركات[22].

ويمكن أن تكون المعلومات المغلوطة والمعلومات المضللة خطيرة بل وقد تكون مميتة، وخاصة في أوقات الأزمات أو الطوارئ أو النزاعات. ففي خلال جائحة مرض فيروس كورونا (كوفيد-19)، غمر الإنترنت طوفان من المعلومات المغلوطة والمعلومات المضللة حول الفيروس وتدابير الصحة العامة واللقاحات[23]. واستغل بعض الأطراف الارتباك الحاصل لتحقيق مآربهم الخاصة، حيث نظم المناهضون للقاحات حملات دفعت المستخدمين إلى مواقع تبيع علاجات مزيفة أو تدابير وقائية[24]. ورفض العديد من ضحايا كوفيد-19 الحصول على التطعيم أو اتخاذ الاحتياطات الصحية الأساسية بعد اطلاعهم على معلومات مغلوطة ومعلومات مضللة على الإنترنت[25].

وكذلك يمكن أن تكون المعلومات المضللة قاتلة في الظروف الاجتماعية والسياسية المتقلبة بالفعل. ففي تقرير صدر عام 2022، درست المقررة الخاصة المعنية بتعزيز وحماية الحق في حرية الرأي والتعبير تأثير المعلومات المستخدمة كسلاح في زرع البلبلة وتغذية الكراهية والتحريض على العنف وإطالة أمد النزاع[26]. وفي تقرير آخر صدر في عام 2022، وُجد أن التضليل الإعلامي يمكن أن ينطوي على التعصب وخطاب الكراهية الموجهين للأقليات والنساء ولأي شخص يُنعت بأنه من "الآخرين"، وهو ما لا يهدد الأشخاص المستهدفين

وجدتْ دراسةٌ بشأن بيانات مستقاة من أجوبة مشاركين في استقصاءات في

142 دولة أن

58,5 في المائة

من الأشخاص الذين يرتادون الإنترنت ووسائل التواصل الاجتماعي بانتظام في **جميع أنحاء العالم** **قلقون** من خطر التعرض

للمعلومات المغلوطة على الإنترنت

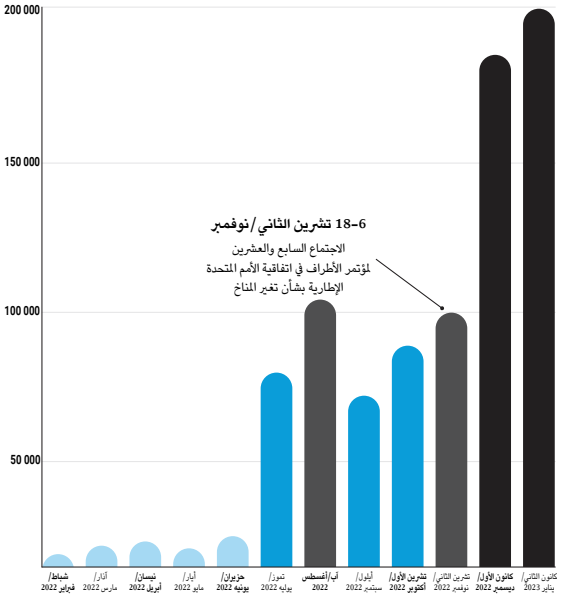
مباشرة وحدهم، بل يهدد أيضا الإدماج والتماسك الاجتماعي. ويمكن أن يفضي ذلك التضليل إلى تضخيم التوترات والانقسامات في حالات الطوارئ أو الأزمات أو اللحظات السياسية الكبرى أو النزاع المسلح[27].

ويقع بعض أسوأ آثار الأضرار على الإنترنت في سياقات لا تكثرت بها المنصات الرقمية، وذلك يشمل حتى السياقات التي تتمتع فيها المنصات بشعبية كبيرة. فالبلدان التي تمر بنزاعات، أو التي تشهد قلاقل وتكون غالبا سوقا لا تدر ربحا وفيرا، لم تخصص لها موارد كافية لضبط المحتوى أو مد المساعدة للمستخدمين. وعلى الرغم من أن وسائل الإعلام التقليدية لا تزال مصدرا مهما للأخبار بالنسبة إلى معظم الناس في مناطق النزاع، فإن الكراهية المستشرية في المنصات الرقمية قد أثارت العنف وأججته أيضا[28]. وواجه بعض المنصات الرقمية انتقادات لدورها في النزاعات، بما في ذلك الحرب المستمرة في أوكرانيا[29].

وكذلك فإن المعلومات المغلوطة والمعلومات المضللة التي تروِّج حول حالة الطوارئ المناخية تؤخِّر الإجراءات التي يلزم اتخاذها على وجه الاستعجال لضمان مستقبل صالح

الشكل الأول

عدد استخدامات الوسم #CLIMATESCAM شهريا على تويتر



المصدر: إدارة التواصل العالمي، باستخدام البيانات المستقاة من برنامج Talkwalker.

للعيش على ظهر الكوكب. ويمكن فهم المعلومات المغلوطة والمعلومات المضللة المتعلقة بالمناخ بكونها المحتوى الخاطئ أو المضلل الذي يقوض الأساس المتفق عليه علميا لظاهرة تغير المناخ الناجم عن النشاط البشري وأسبابه وآثاره. وتسعى حملات منسقة إلى إنكار التوافق العلمي للهيئة الحكومية الدولية المعنية بتغير المناخ أو تحقيره أو صرف الانتباه عنه وعرقلة الإجراءات الملحة التي يتعين اتخاذها لتحقيق أهداف اتفاق باريس لعام 2015. ولا تزال فئة ضئيلة من منكري علم المناخ المجاهرين بآرائهم[30] ترفض الموقف التوافقي وتستأثر بحضور مفرط في بعض المنصات الرقمية. ففي عام 2022، على سبيل المثال، كشفت عمليات محاكاة عشوائية قادتها منظمات تابعة للمجتمع المدني أن خوارزمية فيسبوك تزكي المحتوى الذي ينشره منكرو حقيقة تغير المناخ على حساب علم المناخ[31]. وعلى تويتر، ارتفع عدد استخدامات الوسم #climatescam من أقل من 2 700 استخدام شهريا في النصف الأول من عام 2022 إلى 80 000 استخدام في تموز/يوليه و 199 000 في كانون الثاني/يناير 2023. وأوردت المنصةُ تلك العبارة أيضا باعتبارها أبرز النتائج المتعلقة بالبحث عن كلمة "مناخ"[32]. وفي شباط/فبراير 2022، أدانت الهيئة الحكومية الدولية المعنية بتغير المناخ نشر المعلومات المضللة بشأن تغير المناخ لأول مرة، وأشارت إلى أن "التقويض المتعمد للعلم" يساهم في "إساءة فهم التوافق العلمي ونشر البلبلة وعدم الاكتراث بالمخاطر وبالطابع الاستعجالي والإصرار على معارضة الحقيقة"[33].

عادة ما يلجأ بعض الشركات المنتجة للوقود الأحفوري إلى استراتيجية "التمويه الأخضر"، حيث تُوهِم الجمهور بأن شركة أو كيانا ما يبذل المزيد لحماية البيئة ويقلل من الأنشطة المضرة بها، وهي دعوى غير مطابقة للحقيقة. ولا تتصرف الشركات بمفردها. ذلك أن الكيانات العاملة في مجال الإعلانات والعلاقات العامة وشركات تكنولوجيا الإعلانات والوسائط الإعلامية والمنصات الرقمية، كلُّها تسهل وتدعم الجهود المبذولة لإرباك الجمهور وتحويل الانتباه بعيدا عن مسؤولية قطاع إنتاج الوقود الأحفوري[34].وتكسب شركات الإعلانات والعلاقات العامة التي تنشئ محتوى التمويه الأخضر والأطراف الثالثة التي توزعه بلايين الدولارات بشكل جماعي من تلك الجهود لحماية قطاع إنتاج الوقود الأحفوري من التدقيق والمساءلة. ونظمت شركات العلاقات العامة مئات الحملات لمصلحة شركات الفحم والنفط والغاز[35].

قال 70 في المائة

من حفظة السلام التابعين للأمم المتحدة

إن المعلومات المغلوطة والمعلومات المضللة لها تأثير بالغ الخطورة أو خطير أو معتدل على عملهم

وتؤثر المعلومات المغلوطة والمعلومات المضللة في الديمقراطية تأثيرا عميقا، حيث تُضعف الثقة في المؤسسات الديمقراطية ووسائل الإعلام المستقلة، وتثبِّط المشاركة في الشؤون السياسية والشؤون العامة. ويمكن أن تؤدي إشاعة معلومات كاذبة ومضللة طيلة الدورة الانتخابية إلى حرمان الناخبين من فرصة اتخاذ قرارات مستنيرة بشأن الخيارات المتاحة. ويمكن أن يؤدي انتشار المعلومات المغلوطة والمعلومات المضللة إلى إضعاف ثقة الجمهور في المؤسسات الانتخابية والعملية الانتخابية نفسها – مثل تسجيل الناخبين والاقتراع والنتائج – وقد يؤدي إلى عدم اكتراث الناخبين أو رفض النتائج الانتخابية ذات المصداقية. وقد تبين أن الدول والقادة السياسيين يمكن أن يكون كل منهما مصدرا قويا لإشاعة المعلومات المضللة، حيث يتعمد نشر الأكاذيب بطريقة استراتيجية للاحتفاظ بالسلطة أو تأمينها، أو تقويض العمليات الديمقراطية في بلدان أخرى[36].

وتُستهدف الفئات المهمشة والضعيفة أيضا بالمعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية بشكل متكرر، وهذا يؤدي إلى زيادة إقصائها اجتماعيا واقتصاديا وسياسيا. وتُستهدف المرشحات والناخبات وموظفات الانتخابات والصحفيات وممثلات المجتمع المدني بالمعلومات المضللة المجنسة على الإنترنت[37]. وتقوض هذه الهجماتُ المشاركةَ السياسية وتُضعف المؤسسات الديمقراطية وحقوق الإنسان، بما في ذلك حق هذه الفئات في حرية التعبير والحصول على المعلومات[38]. ويجب أن يولي المجتمع الدول بشكل متزايد أولوية عاجلة لهذه الظواهر، لأسباب أخَصُّها أن أكثر من بليوني ناخب من المقرر أن يتوجهوا إلى صناديق الاقتراع في جميع أنحاء العالم في عام 2024.

كذلك، فإن المعلومات المغلوطة والمعلومات المضللة تُتناقل فيما بين المنصات ووسائل الإعلام التقليدية وداخلها، ويصبح تتبُّعها والتصدي لها أكثر تعقيدا إذا لم تُكتشف في منبعها. ويمكن أن يكون نشر المعلومات المضللة أسلوبا متعمدا تنتهجه وسائل الإعلام المتأثرة إيديولوجيا التي تستحوذ عليها المصالح السياسية ومصالح الشركات[39]. وفي الوقت نفسه، أدى ظهور المنصات الرقمية إلى انخفاض كبير في وسائل الإعلام المستقلة الجديرة بالثقة. وقد رحل متتبعو الأخبار وعائدات الإعلانات بأعداد كبيرة صوب منصات الإنترنت – وهو اتجاه تفاقم بسبب جائحة كوفيد-19. ويمكن معاينة ظاهرة "انقراض وسائل الإعلام" أو "صحارى الأخبار"، حيث تفقد المجتمعات المحلية مصادر الأخبار المحلية الجديرة بالثقة، في بعض المناطق أو البلدان[40]، وهذا ما يساهم في تلوث المنظومة المعلومية. وكثيرا ما لا ينبَّه بشكل كاف على "التمويه الإخباري" – وهو يكون في الحالات التي يُعرَض فيها المحتوى المدعوم بطريقة تجعله يبدو وكأنه نبأ حقيقي – عند نشره على المنصات الرقمية، فيضفي عليه ذلك حُلَّة الشرعية. وبمجرد أن تلتقطه وسائل الإعلام الأخرى، أو يستشهد به السياسيون، أو يتناقله الناس على نطاق واسع عبر المنصات، يصبح أصل المعلومات غامضا بشكل متزايد، ويعجز مستهلكو الأخبار عن تمييز تلك المعلومات عن الحقيقة الموضوعية.

وتؤثر المعلومات المضللة أيضا تأثيرا مباشرا على عمل الأمم المتحدة. وقد أعرب المنسقون المقيمون والمبعوثون والوسطاء وحفظة السلام عن قلقهم من تأثير المعلومات المضللة على السلامة العملياتية للمنظمة وفعاليتها وقدرتها على الإنجاز. ففي استطلاع أجري عام 2022، قال 70 في المائة من حفظة السلام التابعين للأمم المتحدة إن المعلومات المغلوطة والمعلومات المضللة لها تأثير بالغ الخطورة أو خطير أو معتدل على عملهم، بينما قال 75 في المائة إن لها تأثيرا على سلامتهم وأمنهم. وكذلك يمكن استخدام المعلومات المغلوطة والمعلومات المضللة لاستهداف الجهات الفاعلة في مجال العمل الإنساني وإعاقة العمليات المنقذة للحياة في مناطق النزاع.

الشكل الثاني

سلامة المعلومات وأهداف التنمية المستدامة

كما هو مبين، يمكن أن يكون للتهديدات التي تتعرض لها سلامة المعلومات تأثير سلبي على تحقيق أهداف التنمية المستدامة.

ما زالت المعلومات المغلوطة والمعلومات المضللة تقوض الجهود المبذولة في سبيل القضاء على الفقر وتؤثر في الاقتصاد العالمي. ويمكن أن تؤدي الضائقة الاقتصادية أيضا إلى انتشار الأكاذيب الاستقطابية المفعمة بالكراهية، بما في ذلك بشأن الفئات المهمشة. وكانت أزمة تكلفة المعيشة أرضية خصبة بوجه خاص لنشر معلومات مضللة اعتبرت التحولَ إلى استخدام الطاقة المتجددة زوراً سبباً في ارتفاع تكاليف الطاقة أو فقدان الوظائف، على سبيل المثال.

يمكن أن تؤدي الأخطار التي تتهدد نزاهة المعلومات إلى تفاقم الجوع على نطاق العالم، بعوامل من بينها تأجيج النزاعات واستفحال آثار تغير المناخ والكوارث والفقر وعدم المساواة. ويمكن للمعلومات المضللة أن تصرف الانتباه عن التحديات التي تواجه الأمن الغذائي العالمي من جراء النزاعات وتلهي عن الاكتراث بها.

خلال جائحة مرض فيروس كورونا (كوفيد-19)، أدى الوباء المعلوماتي الذي رافقها والذي كان مفعما بالمعلومات المغلوطة والمعلومات المضللة إلى تقويض تدابير الصحة العامة وحملات التطعيم. وما زال الخطرُ الذي يتهدد صحة الأطفال ورفاهيتهم بسبب المعلومات المغلوطة والمضللة المناهضة للتطعيم مستمراً.

يمكن أن يكون للمعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية تأثير سلبي يحول دون إمكانية حصول الناس على التعليم الجيد، وخاصة أفراد الفئات المهمشة، وهم يشملون الشابات والفتيات. وبوسع إمكانية الحصول على المعلومات وحملات تلقين أساسيات استعمال الوسائط الرقمية لزيادة المرونة أن يؤديا دورا رئيسيا في الحد من التأثير المجتمعي لأشكال الأذى المتكبَّد على الإنترنت.

يسعى خطاب الكراهية والمعلومات المضللة القائمين على الفوارق الجنسانية إلى إخضاع النساء بشكل منهجي بإخراسهن وإقصائهن عن المجال العام. ويمكن أن يكون لهما عواقب مدمرة تتراوح بين قمع أصوات النساء وتأجيج الرقابة التي يمارسنها على أنفسهن وبين التسبب لهن في أضرار مهنية وتلطيخ سمعتهن والتحريض على العنف البدني ضدهن.

يعيش بليونان من الناس محرومين من خدمات مياه الشرب المدارة بأمان. ويمكن أن يكون للمعلومات المغلوطة والمعلومات المضللة بشأن سلامة مياه الشرب والصرف الصحي عواقب صحية خطيرة[أ].

إن المعلومات المغلوطة والمعلومات المضللة المتعلقة بالمناخ، التي هي في جزء كبير منها من ترويج قطاع الوقود الأحفوري، تقوض الانتقال العاجل إلى أشكال أنظف لإنتاج الطاقة، وهذا ما يضيق النافذة الآخذة في الانغلاق لتحقيق مستقبل مستدام للجميع.

أشارت الأبحاث إلى الآثار الضارة للمعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية على الاقتصاد[ب].

يمكن أن تؤثر المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وما يُتخذ من تدابير مهلهلة لمواجهة هذه الظواهر تأثيرا مضرا على الابتكار، فتتسبب في أمور من بينها تقليص إمكانات الفئات المهمشة وجعل المساحات الرقمية أقل مساواة وشمولا.

تؤدي المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية المنتشر على الإنترنت إلى استقطاب المجتمعات واستهداف الفئات المهمشة والضعيفة أصلا، ويمكن أن تؤدي إلى زيادة إقصائها اجتماعيا واقتصاديا وسياسيا.

يمكن أن تتقوض الجهود المبذولة من أجل إكساب المدن والمجتمعات المحلية المزيد من مقومات الاستدامة بسبب المعلومات المضللة التي تنكر آثار النشاط البشري على البيئة أو تصرف عنها الانتباه.

استُهدف النشطاء الذين يقفون وراء المبادرات الرامية إلى تعزيز الاقتصاد الدائري وتحفيز الممارسات الهادفة إلى التخلص نهائيا من النفايات بحملة تنشر المعلومات المضللة وخطاب الكراهية ضدهم على الإنترنت.

تؤدي المعلومات المضللة المتعلقة بالمناخ وما تشجع عليه من جمود إلى نسف الجهود المبذولة من أجل اتخاذ إجراءات عاجلة للتصدي لأزمة المناخ، ويشمل ذلك إعاقة التحول الحاسم من الوقود الأحفوري الملوث إلى الطاقة المتجددة وعرقلة الاستثمارات العاجلة في بناء القدرة على مقاومة تغير المناخ.

يمكن أن يكون للمعلومات المغلوطة والمعلومات المضللة تأثير سلبي على الجهود المبذولة للحفاظ على المحيطات والبحار والموارد البحرية واستخدامها على نحو مستدام.

استُهدف النشطاء في مجال البيئة العاملون على حماية الحياة في البر بحملات على الإنترنت تبث الكراهية ضدهم وتروج المعلومات المضللة وتؤدي إلى عواقب واقعية ضدهم. وتقوض المعلومات المغلوطة والمعلومات المضللة المروَّجة بشأن المناخ جهود التصدي للتغير المناخي[ج].

استُخدمت المعلومات المضللة وخطاب الكراهية للتأثير على الانتخابات والخطاب العام وزرع بذور الفتنة. فقد استخدم كل منهما لتقويض الخصوم، وعرقلة جهود صنع السلام، والتحريض على العنف، وإطالة أمد النزاعات، ونسف الثقة في سيادة القانون. وتضررت بشكل خطير من جراء ذلك الجهودُ الرامية إلى تعزيز المجتمعات السلمية والشاملة للجميع وكذا دور الأمم المتحدة في دعم إحلال السلام والأمن[د].

يمكن للمعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية عرقلة الشراكات الجدية في بلوغ الأهداف، في حين يمكن أن يؤدي صرف الموارد إلى معالجة المشكلة إلى إضعاف الجهود الرامية إلى عدم ترك أحد خلف الركب.

أ World Health Organization and United Nations Children's Fund (UNICEF), *Progress on Household Drinking Water, Sanitation and Hygiene 2000–2020: Five Years into the SDGs* (Geneva, 2021).

ب See Roberto Cavazos and CHEQ, "The economic cost of bad actors on the Internet: fake news, 2019"; and London Economics, "The cost of lies: assessing the human and financial impact of COVID-19 related online misinformation on the UK", December 2020.

ج Global Witness, *Last Line of Defence: The Industries Causing the Climate Crisis and Attacks against Land and Environmental Defenders* (2021).

د الوثيقة A/77/288.

الأشهُر المستغرَقَة لإضافة

100 مليون

مستخدم نشط شهريا على الصعيد العالمي

المصدر: موقع SIMILARWEB باستخدام بيانات من موقع SENSOR TOWER

إنستغرام

30 شهرا

تيك توك

9 أشهر

شات جي بي تي

شهران

ما السبيل إلى تعزيز سلامة المعلومات؟

لا تولد المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية من عدم. فهي تنتشر عندما يحس الناس بالإقصاء واللامبالاة أو يعانون من التفاوت الاقتصادي أو يصابون بخيبة أمل سياسية. ولذلك يجب أن تتوخى التدابير المتخذة معالجة هذه التحديات الواقعية. وتعتبر الجهود المبذولة في سبيل تحقيق أهداف التنمية المستدامة شرطا أساسيا لبناء عالم يمكن فيه استعادة الثقة.

ومن المهم لدى وضع التدابير عدم إغفال الفائدة الجليلة التي تعود على العالم بفضل المنصات الرقمية. فهذه المنصات قد أحدثت ثورة في الاتصال الجماهيري في الوقت الفعلي، حيث أتاحت نشر المعلومات المنقذة للحياة أثناء الكوارث الطبيعية والجوائح. وقد ساعدت على حشد الدعم للأهداف التي تسعى الأمم المتحدة إلى تحقيقها، وكثيرا ما أثبتت أنها عوامل إيجابية للإدماج والمشاركة في الحياة العامة. وقد وصلت فيما بين جماعات متباينة جغرافيا من الأشخاص المقصيين، منهم من يعانون حالات صحية نادرة، ومجموعة واسعة من الناشطين الذين يعملون من أجل أن يتحول العالم إلى مكان أفضل.

التدابير التنظيمية

لقد جرى نقاش مطول حول مسألة ما إذا كان من الممكن أو من الضروري تحميل المنصات الرقمية المسؤولية القانونية عن المحتوى الذي تنشره على مواقعها. وفي بعض السياقات، تُستخدم بفعالية القوانين القائمة التي تركز على التشهير والتنمر السيبراني والمضايقة للتصدي للتهديدات التي تتعرض لها سلامة المعلومات، ولكن دون فرض قيود جديدة على حرية التعبير (A/77/287، الفقرة 44).

وعلاوة على ذلك، بُذلت جهود تشريعية في الآونة الأخيرة للتعامل مع المسألة على الصعيدين الإقليمي والوطني. ومن هذه الجهود الإطار التشريعي الذي اعتمده الاتحاد الأوروبي في عام 2022، والذي يشمل قانون الخدمات الرقمية، والمبادرة المتعلقة بالشفافية واستهداف الإعلانات السياسية، و مدونة قواعد الممارسات المتعلقة بالتضليل الإعلامي. ويضع قانون الخدمات الرقمية قواعد جديدة للمستخدمين والمنصات الرقمية والشركات العاملة على الإنترنت داخل الاتحاد الأوروبي. وتستهدف التدابير المتخذة المحتوى والسلع والخدمات غير القانونية على الإنترنت وتوفر آلية للمستخدمين للإبلاغ عن المحتوى غير القانوني والطعن في القرارات التي تخالف مصلحتهم. وهي تقتضي من المنصات الرقمية تحسين الشفافية، لا سيما فيما يتعلق باستخدام خوارزميات التوصية بمنتجات إضافية للمستهلكين وبنوع تلك الخوارزميات، مثلما تقتضي من المنصات الكبرى تزويد الباحثين بإمكانية الاطلاع على البيانات.

وتحدد مدونة قواعد الممارسات المتعلقة بالتضليل الإعلامي مبادئ للمنصات الإلكترونية وقطاع الإعلانات وتفرض على كل منهما التزامات لمكافحة انتشار المعلومات المضللة على الإنترنت في الاتحاد الأوروبي، ووافقت الأطراف الموقعة على المدونة على تنفيذها[41]. ويشمل ذلك التزاما طوعيا بالمساعدة في منع مروجي المعلومات المضللة من الاستفادة من عائدات الإعلانات، سواء من خلال منع نشر الإعلانات التي تحتوي على معلومات مضللة أم من خلال تجنب نشر الإعلانات بحذاء المحتوى الذي يشتمل على معلومات مضللة. ووافقت الأطراف الموقعة على وسْم الإعلانات السياسية بطريقة أوضح، وإدراج معلومات بشأن الجهة الراعية وتكلفة الإعلان وفترة العرض، وإنشاء قواعد بيانات خاصة بالإعلانات السياسية تتيح إمكانية البحث فيها. والتزمت علاوة على ذلك بتبادل المعلومات بشأن سلوكيات التلاعب الخبيث التي يكون الغرض منها هو إشاعة المعلومات المضللة (مثل الحسابات المزيفة، وتضخيم الرأي بواسطة البرامج الآلية، وأشرطة الفيديو التي تنتحل شخصية الغير أو تُزَيّف حقيقة الأشخاص تحقيقا لأغراض خبيثة) والتي يُكشَف عنها في المنصات التابعة لها وتحديث

السياسات وتنفيذها بانتظام من أجل التصدي لتلك الممارسات. وتركز الالتزامات الأخرى على تمكين المستخدمين من التعرف على المعلومات المضللة وفهمها والتبليغ عنها، وعلى تعزيز التعاون مع المحققين في صحة المعلومات وتمكين الباحثين من الحصول على البيانات بطريقة أفضل. ويكمن الاختبار الحقيقي لهذه الآليات الجديدة في تنفيذها.

ويكمن أحد الأهداف الرئيسية لمدونة قواعد الممارسات المتعلقة بالتضليل الإعلامي في تحسين شفافية المنصات. ففي شباط/فبراير 2023، نشرت الأطراف الموقعة على مدونة قواعد الممارسات تقاريرها الأساسية الأولى حول سبل تنفيذ الالتزامات. وقدمت التقارير نظرة متعمقة عن نطاق حرمان مروجي المعلومات المضللة وغيرهم من الجهات التي ضُبط تورطها في سلوك تلاعبي من الاستفادة من عائدات الإعلانات، ومنها جهات ضالعة في جهد منسق واسع النطاق للتلاعب بالرأي العام بشأن الحرب في أوكرانيا في عدة بلدان أوروبية[42].

التدابير المتخذة من جانب المنصات الرقمية

تتميز المنصات الرقمية بشدة تنوعها من حيث الحجم والوظيفة والهيكل وقد اتخذت مجموعة واسعة من التدابير لمعالجة الضرر. والتزمت عدة من المنصات الكبرى علنا بدعم المبادئ التوجيهية بشأن الأعمال التجارية وحقوق الإنسان[43]، ولكن لا تزال ثمة ثغرات في السياسات والشفافية والتنفيذ. ولا تقوم بعض المنصات بإنفاذ معاييرها الخاصة وتسمح، بدرجات متفاوتة، بنشر الأكاذيب والكراهية وتضخيمهما[44]. وتهدف الخوارزميات التي وضعتها المنصات لتعزيز نموذج جني الأرباح إلى الزيادة بشكل متعمد في حجم المشاركة والاستثار بالاهتمام، ويفضي ذلك عادة إلى دفع المستخدمين صوب الاهتمام بالمحتوى الذي يؤدي إلى الاستقطاب أو يكون استفزازيا.

ووضع معظم المنصات الرقمية نوعا ما من أنواع النظم التنظيمية الذاتية أو آليات ضبط المحتوى أو الرقابة، ومع ذلك لا تزال الشفافية بشأن السياسات والممارسات المتعلقة بإزالة المحتوى تشكل تحديا[45]. ويتسم الاستثمار في هذه الآليات عبر المناطق واللغات بتفاوته الشديد ويتمركز إلى حد كبير في بلدان الشمال، وكذلك هو الشأن بالنسبة إلى إنفاذ المنصات لقواعدها الخاصة. ووجدت دراسة استقصائية حديثة[46] أن ترجمة برامج التحكم في المحتوى وآليات الرقابة إلى اللغات المحلية ناقصة في جميع المنصات. وفي الوقت نفسه، كثيرا ما يستعان بمصادر خارجية لضبط المحتوى ويعاني هذا النشاط في اللغات الأخرى غير الإنكليزية من نقص ذريع في الموارد[47]. وأثارت شهادات القيِّمين على ضبط المحتوى أسئلة مقلقة تتعلق بسوء المعاملة ومعايير العمل والصدمات التبَعِيّة[48]. ويبلغ القيمون على ضبط المحتوى عن تعرضهم المستمر لمحتوى عنيف ومزعج، ويُمنَحون بضع ثوان لتحديد ما إذا كان المحتوى المبلغ عنه ينتهك سياسات الشركة. ويمكن أن تؤدي النظم الآلية لضبط المحتوى دورا أساسيا، ولكنها معرضة للتحيز المحتمل بناء على البيانات والهياكل المستخدمة لتدريبها. وهي تسجل أيضا معدلات عالية من حيث ما يشوبها من أخطاء في اللغة الإنكليزية، إلا أن ما تقعُ فيه من أخطاء في سائر اللغات أفدحُ. ويوظِّف عدد من المنصات الرقمية أفرقة معنية بمسائل الثقة والأمان وحقوق الإنسان وسلامة المعلومات، ومع ذلك كثيرا ما لا يُشرَك هؤلاء الخبراء في المراحل الأولى من تصميم المنتج كما أنهم في حالات كثيرة يكونون أول من يُستغنى عن وظائفهم عند اتخاذ تدابير لتخفيض التكاليف.

إمكانية الحصول على البيانات

تمكين الباحثين من الحصول على البيانات هي إحدى الأولويات الملحة على النطاق العالمي. ولا تزال البحوث والموارد القائمة موجهة بشكل كبير للولايات المتحدة الأمريكية وأوروبا. ولم ينشر سوى عدد محدود من الأبحاث بشأن وقع ذلك على بقية العالم، ما عدا بعض الاستثناءات الملحوظة، منها التقارير المتعلقة بعمليات الأمم المتحدة لحفظ السلام في أفريقيا والبعثة الدولية المستقلة لتقصي الحقائق في ميانمار[49]، وبعض التقارير الاستقصائية والتغطية التي يُعِدها الصحفيون[50]. ويرجع ذلك في جانب منه إلى أن الباحثين تُعوزُهم إمكانية الوصول إلى المنصات والاطلاع على بياناتها. كذلك، كثيرا ما تكون الأدوات اللازمة للبحث الفعال في البيانات المحدودة التي توفرها المنصات مصممة على نحو يراعي اعتبارات التسويق وهي بالإضافة إلى ذلك باهظة الثمن بدرجة كبيرة. ومن شأن تحول المنصات من العمل بنهج "الحصول على البيانات بناء على الطلب" إلى الأخذ بنهج "الكشف التلقائي عن البيانات"، مع وجود الضمانات اللازمة لصون الخصوصية، أن يتيح للباحثين تقييم الأضرار بشكل صحيح.

تمكين المستخدمين

أجرت هيئات المجتمع المدني والأكاديميون أبحاثا مكثفة حول أفضل السبل للتصدي لمشكلة نشر المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية والحرص في الوقت ذاته على حماية حرية التعبير. وشدد عدد منهم على وضع حلول تنطلق من القاعدة بحيث تمكن مستخدمي الإنترنت من تقليص تأثير الأضرار الناجمة عن استعمال الإنترنت على مجتمعاتهم المحلية وتَنزَع من أيدي المنصات مركزيةَ السلطة.

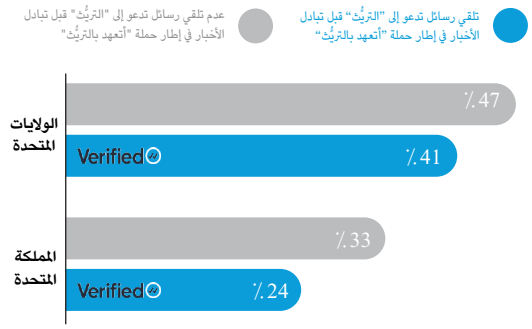
وينبغي تشجيع مستخدمي المنصات، ومنهم أفراد الفئات المهمشة، على الإقبال على الحيز السياساتي وإدماجهم وإشراكهم فيه. ويمتلك الشباب خاصة خبرات زاخرة ومتعمقة. وباعتبار نشأتهم في العصر الرقمي، فإن الشباب، وبخاصة الإناث منهم، والأطفال كثيرا ما يكونون بالفعل هدفا للمعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وسوف يتأثرون بالمنصات الناشئة والجديدة تأثرا مباشرا. ويمكن للمستخدمين الأصغر سنا أن يتحدثوا انطلاقا من تجاربهم عن التأثير المتمايز للمقترحات المختلفة وعيوبها المكنة. وقد ساهموا بهِمة في جهود الدعوة والتحقق من صحة المعلومات على الإنترنت[51].

ويمكن أن يساعد صقل مهارات التفكير النقدي لدى رواد الإنترنت على أن يصبحوا أقدر على الصمود في وجه ممارسات التلاعب الرقمي. وعلى وجه التحديد، فإن الإلمام بأبجديات

الشكل الثالث

فعالية حملات الأمم المتحدة في مكافحة المعلومات المغلوطة والمعلومات المضللة

احتمال تبادل أخبار زائفة (2021)[أ]



أ استنادا إلى بحث أجراه معهد ماساتشوستس للتكنولوجيا في آذار/مارس 2021.

التكنولوجيا الرقمية يُعلِّم المستخدمين تقييم المعلومات التي يصادفونها على في الإنترنت بشكل أفضل ونقلها بطريقة مسؤولة. وقد اكتسبت مجموعة من كيانات الأمم المتحدة تجارب قيمة في هذا الميدان. فقد نجحت مبادرة الأمم المتحدة التي تحمل شعار "أتعهد بالتريث"[52] في استعمال عدد من الأساليب، من بينها الرسائل الموجهة للمستخدمين، وفضح الأكاذيب مسبقا بتحذير المستخدمين منها قبل أن يصادفونها، وحملات التشجيع على الإلمام بأبجديات التكنولوجيا الرقمية.

إزالة الحوافز

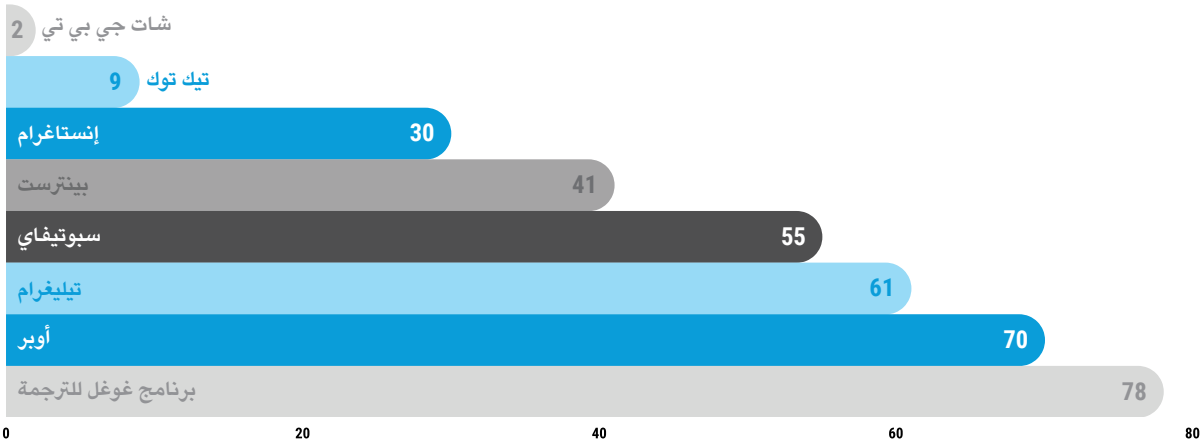
نماذج الأعمال الحالية لمعظم المنصات الرقمية تعطي الأولوية للتفاعلُ على حساب حقوق الإنسان والخصوصية والسلامة. ويشمل ذلك تحويل البيانات الشخصية إلى قيمة نقدية من أجل تحقيق الربح، على الرغم من الأدلة المتزايدة على الضرر المجتمعي الناجم عن نموذج الأعمال هذا.

واستكشف بعض مجموعات المجتمع المدني والباحثين سبل منع الجهات التي تختلق المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وتنشرها على الإنترنت من الاستفادة من عائدات الإعلانات ومن ثم تجريدها من الحوافز التي تشجعها على القيام بذلك، وأشاروا إلى أن حرية التعبير إذا كانت حقا أساسيا من حقوق الإنسان، فإن استغلالها لكسب الربح ليس كذلك[53]. وتسعى مقترحاتهم إلى التصدي للمكاسب المالية المجنية من وراء إشاعة المعلومات المضللة، وضمان الشفافية الكاملة حول الإيرادات المالية المتأتية من نشر المحتوى والتقييمات المستقلة للمخاطر، وتثبيط المشاركين في الإعلان على الإنترنت عن السماح بنشر المعلومات المضللة.

والعلامات التجارية التي تنشر إعلاناتها جنبًا إلى جنب مع المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية إنما تخاطر بتقويض فعالية حملاتها والإساءة إلى سمعتها في نهاية المطاف. ويمكن للمعلنين وضع سياسات واضحة لتجنب القيام بشكل غير مقصود بتمويل المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وإضفاء الشرعية عليها، وللمساعدة في جعلها غير مربحة. ويمكن أن تشمل إجراءات التنفيذ إدارة قوائم محدثة للإدراج والاستبعاد واستخدام أدوات التحقق من الإعلانات. ويمكن للمعلنين أيضًا الضغط على المنصات الرقمية لتكثيف الإجراءات الرامية إلى حماية سلامة المعلومات، ويمكنهم الامتناع عن الإعلان لدى وسائل الإعلام التي تغذي الكراهية وتنشر المعلومات المضللة[54].

الشكل الرابع

الأشهُر المستغرَقة لإضافة 100 مليون مستخدم نشط شهريا لبرنامج شات جي بي تي مقارنة بالتطبيقات الشائعة الأخرى



المصدر: موقع Similarweb باستخدام بيانات من موقع Sensor Tower.

وسائل الإعلام المستقلة

ما زالت التدابير الجديدة التي تتخذ في عشرات البلدان تقوض حرية الصحافة. فحسب تقرير اليونسكو العالمي لعام 2022 الصادر ضمن سلسلتها الرائدة عن الاتجاهات العالمية في مجال حرية التعبير وتطور وسائط الإعلام، شهد 85 في المائة من سكان العالم انخفاضا في حرية الصحافة في بلدانهم خلال السنوات الخمس السابقة[55]. ومع استمرار وجود 2,7 بليون شخص غير قادرين على الاتصال بالإنترنت[56]، فإن من الأولويات الإضافية العمل على تعزيز وسائل الإعلام المستقلة، وتحفيز انتشار مبادرات التحقق من صحة المعلومات، ودعم التقارير الموثوقة والدقيقة من أجل المصلحة العامة. فالنقاش العام الحقيقي يعتمد على الحقائق التي تروى بوضوح ويبلغ عنها باستقلالية وحس أخلاقي. إن المندوبين الصحفيين ذوي الحس الأخلاقي والتدريب الجيد والعاملين في ظروف عمل جيدة يمتلكون ما يلزم من المهارات لاستعادة التوازن في وجه انتشار المعلومات المغلوطة والمعلومات المضللة. ويمكنهم إسداء خدمة حيوية ألا وهي توفير معلومات دقيقة وموضوعية وموثوقة حول القضايا المهمة.

التحوط للمستقبل

يجب علينا، حتى ونحن نبحث عن حلول لحماية سلامة المعلومات في الظروف الحالية، ضمان أن تحافظ التوصيات على جدواها في المستقبل، بحيث تعالج المسائل الناجمة عن التكنولوجيات الناشئة وتلك التي لم تظهر بعد. وقد أُعلن عن إنشاء منصة شات جي بي تي (ChatGPT-3) التابعة لشركة Open AI للذكاء الاصطناعي في تشرين الثاني/نوفمبر 2022، وبلغ عدد مستخدميها 100 مليون مستخدم بحلول كانون الثاني/يناير 2023، وهذا ما يجعلها أسرع التطبيقات الجماهيرية نموا في التاريخ[57]، مع تسابق العديد من الشركات الأخرى من أجل تطوير أدوات منافسة. إن تكنولوجيات الذكاء الاصطناعي وإن كانت تزخر بقدرة يكاد يتعذر تصورها في المساعدة على مواجهة التحديات العالمية إلا أن ثمة مخاوف حقيقية وملحة من القدرة القوية المماثلة للتطورات الحديثة التي شهدتها تلك التكنولوجيات - بما في ذلك برمجيات توليد الصور وأشرطة الفيديو المزيفة - على تهديد سلامة المعلومات. فقد أظهرت التقارير والأبحاث الحديثة أن أدوات الذكاء الاصطناعي التوليدية أنتجت معلومات مغلوطة ومعلومات مضللة وخطاب كراهية، وقُدمت للمستخدمين بطريقة مقْنِعة على أنها معلومات حقيقية[58].

ويقود مبعوثي المعني بالتكنولوجيا الجهود الرامية إلى تقييم الآثار المترتبة على الذكاء الاصطناعي التوليدي والمنصات الناشئة الأخرى. ويجب علينا ونحن نقوم بذلك أن نتعلم من أخطاء الماضي. فقد بدأ استعمال المنصات الرقمية في العالم دون وعي أو تقييم كاف لما يحتمل أن ينجم عنها من أضرار على المجتمعات والأفراد. وها هي تتاح لنا الآن الفرصة لضمان ألا يعيد التاريخ نفسه مع التكنولوجيا الناشئة. ويجب إنهاء عصر فلسفة وادي السيليكون القائمة على شعار "السرعة والاختبار في العمل والابتكار". ومن الضروري العمل على جعل خصوصية المستخدم والأمن والشفافية والأمان عناصر متأصلة في جميع التكنولوجيات والمنتجات الجديدة منذ بداية تصميمها.

التدابير المتخذة من جانب الأمم المتحدة

يجري أيضا اتخاذ خطوات من قِبل جهات تشمل عمليات الأمم المتحدة ومكاتبها القطرية وتهدف إلى رصد خطر المعلومات المغلوطة والمعلومات المضللة على إنجاز ولايات الأمم المتحدة وتحليل ذلك الخطر والتصدي له. وتضع استراتيجية الأمم المتحدة وخطة عملها بشأن خطاب الكراهية التوجيه الاستراتيجي لمنظومة الأمم المتحدة الهادف إلى التصدي لخطاب الكراهية على الصعيدين الوطني والعالمي. وفي شباط/فبراير 2023، استضافت اليونسكو مؤتمر الإنترنت لبناء الثقة لمناقشة مجموعة من مشاريع المبادئ التوجيهية العالمية لوضع ضوابط تنظيمية للمنصات الرقمية، ومن المزمع استكمال إعدادها في وقت لاحق من هذا العام[59].

ومن شأن هذه المبادرات والنهج مجتمعةً أن تحدد طريق المضي قدما في إعداد المبادئ الأساسية لمدونة الأمم المتحدة لقواعد السلوك.

عدد استخدامات الوسم

#CLIMATESCAM

على تويتر شهريا

200

شباط/فبراير 2022

100 ألف

تشرين الثاني/نوفمبر 2022

حصول ارتفاع بتزامن مع
انعقاد الاجتماع السابع والعشرين
لمؤتمر الأطراف في اتفاقية الأمم المتحدة
الإطارية بشأن تغير المناخ

بلغت المشاركات ذروتها
في كانون الثاني/يناير 2023 حيث بلغت

199 300

نحو مدونة للأمم المتحدة لقواعد السلوك

تستند مدونة الأمم المتحدة لقواعد السلوك بشأن سلامة المعلومات على المنصات الرقمية، التي سأطرحها، إلى المبادئ التالية:

- الالتزام بسلامة المعلومات
- احترام حقوق الإنسان
- دعم وسائل الإعلام المستقلة
- زيادة الشفافية
- تمكين المستخدِمين
- تعزيز الأبحاث وإمكانية الحصول على البيانات
- تعزيز التدابير
- تشديد إجراءات التثبيط
- تعزيز الثقة والأمان

وقد استخلصت هذه المبادئ من الأفكار الأساسية التي نوقشت في هذا الموجز السياساتي، وهي منسجمة والموجزَ السياساتي الذي أعددتُه بشأن الاتفاق الرقمي العالمي ومترابطة معه. وستُدعى الدول الأعضاء إلى تنفيذ مدونة قواعد السلوك على الصعيد الوطني. وستتواصل المشاورات مع الأطراف صاحبة المصلحة لمواصلة صقل محتوى مدونة قواعد السلوك وتحديد منهجيات ملموسة لتفعيل مبادئها.

ويمكن أن تستند مدونة قواعد السلوك إلى التوصيات التالية:

الالتزام بسلامة المعلومات

(أ) يجب على جميع الأطراف صاحبة المصلحة الامتناع عن استخدام المعلومات المضللة وخطاب الكراهية أو دعمهما أو تضخيمهما لأي غرض كان، بما في ذلك السعي إلى تحقيق مآرب سياسية أو عسكرية أو أهداف استراتيجية أخرى، أو التحريض على العنف، أو تقويض العمليات الديمقراطية، أو استهداف السكان المدنيين أو من يكونون في حالة ضعف من الفئات أو المجتمعات أو الأفراد؛

احترام حقوق الإنسان

(ب) ينبغي للدول أن تقوم بما يلي:

'1' ضمان أن تكون التدابير المتخذة للتصدي للمعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية متسقة مع القانون الدولي، بما في ذلك القانون الدولي لحقوق الإنسان، وألاّ يساء استخدامُها لمنع أي تعبير مشروع عن وجهات النظر أو الرأي، بما في ذلك من خلال إجراءات التعطيل الشامل لإمكانية الاتصال بالإنترنت أو حظر المنصات أو وسائل الإعلام؛

'2' اتخاذ تدابير تنظيمية لحماية الحقوق الأساسية لمستخدمي المنصات الرقمية، بما في ذلك آليات الإنفاذ، مع الحرص على الشفافية الكاملة فيما يتعلق بالشروط المفروضة على شركات التكنولوجيا؛

(ج) ينبغي لجميع الأطراف صاحبة المصلحة الامتثال للمبادئ التوجيهية بشأن الأعمال التجارية وحقوق الإنسان؛

دعم وسائل الإعلام المستقلة

(د) ينبغي للدول الأعضاء أن تضمن وجود مشهد إعلامي حر وفعال ومستقل وتعددي يوفر حماية قوية للصحفيين ووسائل الإعلام المستقلة، وأن تدعم إنشاء هيئات مستقلة معنية بالتحقق من صحة المعلومات باللغات المحلية وتمويل تلك الهيئات وتدريب أفرادها؛

(هـ) ينبغي لوسائل الإعلام الإخبارية أن تكفل تقارير مستقلة دقيقة وملتزمة بالمبادئ الأخلاقية يقدمها أفراد يُزوَّدون بالتدريب الجيد ويعملون في ظروف ملائمة انسجاما مع القواعد والمعايير الدولية للعمل وحقوق الإنسان؛

تعزيز الشفافية

(و) ينبغي للمنصات الرقمية القيام بما يلي:

'1' ضمان قدرٍ مجدٍ من الشفافية فيما يتعلق بالخوارزميات والبيانات وضبط المحتوى والإعلانات؛

'2' سن سياسات يسهل الاطلاع عليها بشأن المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وتعميمها، والإبلاغ عن مدى انتشار المعلومات المضللة المنسقة فيما تقدمه من خدمات وعن فعالية السياسات الموضوعة لمواجهة ذلك؛

(ز) ينبغي أن تضمن وسائل الإعلام الإخبارية قدرا مجديا من الشفافية فيما يتعلق بمصادر التمويل والسياسات الإعلانية، وأن تميز بوضوح المحتوى التحريري عن الإعلانات المدفوع عنها، بما في ذلك عند النشر على المنصات الرقمية؛

تمكين المستخدمين

(ح) ينبغي للدول الأعضاء ضمان حصول الجمهور على معلومات حكومية دقيقة وشفافة وذات مصادر موثوقة، ولا سيما المعلومات التي تخدم المصلحة العامة، بما في ذلك جميع جوانب أهداف التنمية المستدامة؛

(ط) ينبغي أن تضمن المنصات الرقمية تمكين المستخدم وحمايته في جو من الشفافية وأن تتيح للأشخاص خيارات أوسع بشأن المحتوى الذي يطلعون عليه والكيفية التي تستخدم بها بياناتهم. وينبغي أن تمكِّن المستخدمين من إثبات هويتهم وصحة بياناتهم الشخصية دون أن يُضطروا إلى دفع مقابل نقدي أو التنازل عن الخصوصية لقاء ذلك وأن تضع إجراءات شفافة تتيح لهم أن يقدموا شكاواهم وبلاغاتهم وتكون تلك الإجراءات مدعومة بآليات لاستعراض الشكاوى تتسم بالاستقلالية وتكون معروفة على نطاق واسع ويسهل الوصول إليها؛

(ي) ينبغي لجميع الأطراف صاحبة المصلحة الاستثمار في حملات قوية للتشجيع على الإلمام بأبجديات التكنولوجيا الرقمية لتمكين المستخدمين من جميع الأعمار من فهم كيفية عمل المنصات الرقمية بشكل أفضل، والأوجه الممكنة لاستخدام بياناتهم الشخصية، وتعرُّف المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية والتصدي لها. وينبغي إيلاء اهتمام خاص لضمان أن يكون الشباب والمراهقون والأطفال على وعي تام بحقوقهم في المساحات الإلكترونية؛

تعزيز الأبحاث وإمكانية الحصول على البيانات

(ك) ينبغي للدول الأعضاء أن تستثمر في الأبحاث المستقلة بشأن انتشار المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وتأثيرها عبر جميع البلدان واللغات، لا سيما في البيئات التي تعاني من نقص الخدمات وفي اللغات الأخرى غير الإنكليزية، وإسداء الدعم لتلك الأبحاث وذلك ما سيتيح للمجتمع المدني والأوساط الأكاديمية العمل بحرية وأمان؛

(ل) ينبغي للمنصات الرقمية القيام بما يلي:

'1' السماح للباحثين والأكاديميين بالاطلاع على البيانات، مع احترام خصوصية المستخدمين. وينبغي تمكين الباحثين من جمع أمثلة وبيانات نوعية عن الأفراد والفئات المستهدفة بالمعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية لفهم نطاق الأضرار وطبيعتها بشكل أفضل، مع الحرص على احترام مبدأ حماية البيانات وحقوق الإنسان؛

'2' ضمان المشاركة الكاملة للمجتمع المدني في الجهود الرامية إلى التصدي للمعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية؛

تعزيز التدابير

(م) ينبغي لجميع الأطراف صاحبة المصلحة القيام بما يلي:

'1' تخصيص الموارد اللازمة للتصدي لمنابع المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وانتشارها وتأثيرها والإبلاغ عن تلك المنابع، مع الحرص على احترام قواعد حقوق الإنسان ومعاييرها وزيادة الاستثمار في بناء قدرات التحقق من صحة المعلومات في جميع البلدان والبيئات؛

’2‘ إقامة ائتلافات واسعة للنهوض بسلامة المعلومات، وذلك من خلال الجمع بين الخبرات والنهج المختلفة للمساعدة في سد الفجوة بين المنظمات المحلية وشركات التكنولوجيا العاملة على نطاق عالمي؛

’3‘ تعزيز التدريب وبناء القدرات لتطوير فهم كيفية ظهور المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وتدعيم استراتيجيات منعها والتخفيف من حدتها؛

تشديد إجراءات التثبيط

(ن) ينبغي أن تتخلى المنصات الرقمية عن أسلوب العمل الذي يعطي الأولوية لمدى تفاعل المستخدمين مع المحتوى المنشور على حساب الاعتبارات المتعلقة بحقوق الإنسان والخصوصية والأمان؛

(س) ينبغي للجهات المعلِنة والمنصات الرقمية ضمان عدم نشر الإعلانات بحذاء المعلومات المغلوطة أو المعلومات المضللة أو خطاب الكراهية على الإنترنت، وعدم الترويج للإعلانات التي تحتوي على معلومات مضللة؛

(ع) ينبغي لوسائل الإعلام الإخبارية أن تضمن وَسْم جميع الإعلانات المدفوع عنها والمحتوى الإعلاني بصفتها هذه وسماً واضحا وخُلوَّ تلك الإعلانات وذلك المحتوى من المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية؛

تعزيز الثقة والأمان

(ف) ينبغي للمنصات الرقمية القيام بما يلي:

’1‘ مراعاة اعتبارات الأمان والخصوصية بجعلها عنصرا متأصلا في تصميم جميع المنتجات، بما في ذلك من خلال توفير الموارد الكافية لتوظِّفَ المنصاتُ في هياكلها الداخلية خبراء متخصصين في مسائل الثقة والأمان، مع الحرص إضافة إلى ذلك على تطبيق السياسات في جميع البلدان واللغات على نحو متسق؛

’2‘ الاستثمار في نظم ضبط المحتوى سواء القائمة منها على تقنيات الذكاء الاصطناعي أم التي يشرف عليها البشر بجميع اللغات المستخدمة في البلدان التي تنشط فيها المنصات الرقمية، وضمان شفافية آليات الإبلاغ عن المحتوى، مع تسريع معدل الاستجابة، وبخاصة في حالات النزاع؛

(ص) ينبغي لجميع الأطراف صاحبة المصلحة أن تتخذ تدابير عاجلة وفورية لضمان استخدام تكنولوجيا الذكاء الاصطناعي استخداما مسؤولا يرعى المبادئ الأخلاقية ويكفل الأمان والأمن ويَمْتثل معاييرَ حقوق الإنسان وأن تتصدى لما يترتب على التطورات الأخيرة الحاصلة في هذا الميدان من تداعيات على انتشار المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية.

الخطوات المقبلة

- ستجري الأمانة العامة للأمم المتحدة مشاورات واسعة النطاق مع طائفة من الأطراف صاحبة المصلحة بشأن إعداد مدونة الأمم المتحدة لقواعد السلوك، بما في ذلك وضع آليات للمتابعة والتنفيذ. ويمكن أن يشمل ذلك إنشاء مرصد مستقل يضم خبراء معترفا بهم لتقييم التدابير التي تتخذها الأطراف الملتزمة بمدونة قواعد السلوك، وغيرها من آليات الإبلاغ.
- ويمكن للأمانة العامة للأمم المتحدة، دعماً منها لتنفيذ المدونة وتوجيهه، أن تجري دراسات متعمقة لتحسين فهم موضوع سلامة المعلومات على الصعيد العالمي، ولا سيما في أجزاء العالم غير المشمولة بقدر كاف من الأبحاث.
- وسينشئ الأمين العام آلية في الأمانة العامة للأمم المتحدة تكون مكرسة لتوسيع نطاق التدابير المتخذة للتصدي للمعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية التي تنشر على الإنترنت وتؤثر في تنفيذ ولايات الأمم المتحدة وأولوياتها الموضوعية. وستقوم تلك الآلية، بالاستناد إلى عمليات الرصد والتحليل التي يضطلع بها الخبراء، بوضع استراتيجيات تواصلية تصمَّم خصيصا لاستباق التهديدات و/أو التصدي لها بسرعة قبل أن تتحول إلى ضرر على الإنترنت وخارجها، ودعم بناء قدرات موظفي الأمم المتحدة والدول الأعضاء. وسوف تدعم ما تبذله الدول الأعضاء والمنصات الرقمية والأطراف الأخرى صاحبة المصلحة من جهود التزاماً بالمدونة وتنفيذا لقواعدها بعد الفراغ من إعدادها.

خاتمة

يعتبر تعزيز سلامة المعلومات على المنصات الرقمية أولوية ملحة للمجتمع الدولي. وستكفل التدابير التي تحد من تأثير المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية تعزيز الجهود الرامية إلى تحقيق مستقبل مستدام لا يُترك فيه أحد خلف الركب، سواء تعلق الأمر بالصحة أم المساواة بين الجنسين أم السلام أم العدالة أم التعليم أم إجراءات مكافحة تغير المناخ. وحتى لو اتُّخذت إجراءات على الصعيد الوطني، فإن التصدي لتلك المشاكل بشكل كامل لن يتأتى إلا من خلال تعاون عالمي أمتن. وتبيِّن الأفكار الأساسية المبسوطة في هذا الموجز السياساتي أن الإجراءات المتخذة من أجل تعزيز سلامة المعلومات يجب أن تكون قائمة على رَعْي حقوق الإنسان وعلى تعدد الأطراف المشاركة في وضعها وتعدد الأبعاد التي تتناولها. وقد جرى تصنيفها في عدد من المبادئ التي يتعين النظر فيها من أجل إعداد مدونة للأمم المتحدة لقواعد السلوك للنهوض بسلامة المعلومات على المنصات الرقمية يكون من شأنها أن توفر مخططا لتعزيز سلامة المعلومات في ظل التشبث القوي بتدعيم حقوق الإنسان. وإنني أتطلع إلى العمل مع الدول الأعضاء وغيرها من الأطراف صاحبة المصلحة من أجل ترجمة هذه المبادئ إلى التزامات ملموسة.

المرفق

المشاورات مع الدول الأعضاء ومع الأطراف المعنية الأخرى ذات المصلحة

تستند الأفكار الواردة في هذا الموجز السياساتي إلى المقترحات الواردة في التقرير المعنون "خطتنا المشتركة" (A/75/982)، التي استفادت من مشاورات مكثفة أجريت مع الدول الأعضاء، ومنظومة الأمم المتحدة، وقادة الفكر، والشباب، والجهات الفاعلة في المجتمع المدني من جميع أنحاء العالم. ويستجيب الموجز السياساتي، على وجه الخصوص، للأفكار الغنية والمفصلة التي أدلى بها الدول الأعضاء وغيرها من الأطراف صاحبة المصلحة بشأن خطتنا المشتركة في أثناء المناقشات الخمس والعشرين التي عقدتها الجمعية العامة.

وقبل نشر هذا الموجز السياساتي، أُجريت مشاورات مع الدول الأعضاء، شملت عقد جلسة إحاطة غير رسمية للجنة الإعلام، دُعي إلى حضورها جميع الأطراف غير الأعضاء في اللجنة. وكذلك فقد أُجريت مناقشات مع الجهات الشريكة من هيئات المجتمع المدني والأكاديميين والخبراء والقطاع الخاص، بما في ذلك شركات التكنولوجيا.

وسوف تُجرى مشاورات عريضة القاعدة بشأن إعداد مدونة قواعد السلوك، قبل انعقاد مؤتمر قمة المعني بالمستقبل.

الحواشي

1 A/HRC/42/50; A/77/287; A/HRC/51/53; United Nations, "Statement by Alice Wairimu Nderitu, Special Adviser on the Prevention of Genocide, condemning the recent escalation of fighting in Ethiopia", press release, 19 October 2022; Office of the United Nations High Commissioner for Human Rights (OHCHR), "Myanmar: Social media companies must stand up to junta's online terror campaign say UN experts", press release, 13 March 2023; OHCHR, "Freedom of speech is not freedom to spread racial hatred on social media: UN experts", statement, 6 January 2023; Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, "#JournalistsToo: women journalists speak out", 24 November 2021; and OHCHR, "Sri Lanka: Experts dismayed by regressive steps, call for renewed UN scrutiny and efforts to ensure accountability", press release, 5 February 2021.

2 A/HRC/47/25، الفقرة 15.

3 Kalina Bontcheva and Julie Posetti, eds., *Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression – Broadband Commission Research Report on "Freedom of Expression and Addressing Disinformation on the Internet"* (Geneva, International Telecommunication Union (ITU); Paris, UNESCO, 2020).

4 انظر "Countering disinformation"، United Nations، يمكن الاطلاع عليه عبر الرابط التالي: www.un.org/en/countering-disinformation، والوثيقة A/77/287.

5 يمكن الاطلاع عليه عبر الرابط التالي: https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_EN.pdf.

6 تجري الأمم المتحدة حاليا دراسة لبحث أوجه الترابط والصلة بين المعلومات المغلوطة والمعلومات المضللة وخطاب الكراهية وتحديد المواطن التي يوجد فيها تقارب أو تباعد بين هذه الظواهر المترابطة ولكن المتمايزة على المستويين المفاهيمي والعملي.

7 تعرف المفوضية الأوروبية المنصات الإلكترونية بكونها في "Shaping Europe's digital future: online platforms"، 7 حزيران/يونيه 2022. يمكن الاطلاع عليه عبر الرابط التالي: https://digital-strategy.ec.europa.eu/en/policies/online-platforms.

8 انظر OHCHR, "Moderating online content: fighting harm or silencing dissent?", 23 July 2021.

9 انظر "Countering disinformation"، United Nations، والوثيقة A/77/287.

10 A/HRC/47/25.

11 Stephanie Kirchgaessner and others, "Revealed: the hacking and disinformation team meddling in elections", *The Guardian*, 14 February 2023.

12 Alexandre Alaphilippe and others, "Doppelganger – media clones serving Russian propaganda", EU DisinfoLab, 27 September 2022.

13 United Nations Economist Network, "New economics for sustainable development: attention economy".

14 Twitter, "About Twitter Blue"; and Meta, "Testing Meta Verified to help creators establish their presence", 17 March 2023.

15 A/77/287.

16 في شهر شباط/فبراير 2023، بلغ عدد الدول الأعضاء الأطراف في العهد الدولي الخاص بالحقوق المدنية والسياسية 173 دولة.

17 يجب أن تستوفي القيود المفروضة على حرية التعبير الشروط الراسخة التالية: الشرعية، أي أن القيود يجب أن ينص عليها القانون بطريقة تميز بما يكفي من الدقة بين التعبير المشروع والتعبير غير المشروع؛ والضرورة والتناسب، أي أن يَفرِض التقييد بشكل واضح أقل عبء على ممارسة الحق وأن يحمي بالفعل، أو يرجح أن يحمي، المصلحة المعنية المشروعة للدولة؛ والمشروعية، أي أن القيود، لكي تكون قانونية، يجب ألا تحمي إلا المصالح المذكورة في المادة 19 (3) من العهد الدولي الخاص بالحقوق المدنية والسياسية.

18 تشمل هذه السياقات الاتصالات السمعية البصرية في تونس وكوت ديفوار والمغرب، وعملية رصد التحريض على العنف التي قامت بها بعثة الأمم المتحدة المتكاملة المتعددة الأبعاد لتحقيق الاستقرار في جمهورية أفريقيا الوسطى.

19 Aleksi Knuutila, Lisa-Maria Neudert and Philip N. Howard, "Who is afraid of fake news? Modeling risk perceptions of misinformation in 142 countries", *Harvard Kennedy School (HKS) Misinformation Review*, vol. 3, No. 3 (April 2022).

20 ITU, *Measuring the Information Society* (Geneva, 2013).

21 United Nations Children's Fund (UNICEF), "Protecting children online", 23 June 2022. يمكن الاطلاع عليه عبر الرابط التالي: www.unicef.org/protection/violence-against-children-online.

22 UNESCO, working papers on digital governance and the challenges for trust and safety. يمكن الاطلاع عليه عبر الرابط التالي: www.unesco.org/en/internet-conference/working-papers.

23 انظر Julie Posetti and Kalina Bontcheva, "Disinfodemic: deciphering COVID-19 disinformation", policy brief 1, (Paris, UNESCO, 2020), and "Disinfodemic: dissecting responses to COVID-19 disinformation", policy brief 2 (Paris, UNESCO, 2020).

24 Center for Countering Digital Hate, *Pandemic Profiteers: The Business of Anti-Vaxx* (2021).

25 Michael A Gisondi and others, "A deadly infodemic: social media and the power of COVID-19 misinformation", *Journal of Medical Internet Research*, vol. 24, No. 2 (February 2022).

26 A/77/288.

27 A/77/287، الفقرة 6.

28 في عام 2018، أعلنت بعثة دولية مستقلة لتقصي الحقائق عينها مجلس حقوق الإنسان أن فيسبوك هو " المنصة الرئيسية في نشر خطاب الكراهية في ميانمار" (A/HRC/42/50، الفقرة 72).

29 انظر United Nations News, "Hate speech: a growing, international threat", 28 January 2023, and "Digital technology, social media fuelling hate speech like never before, warns UN expert", 20 October 2022.

30 انظر John Cook, "Deconstructing climate science denial", in *Research Handbook in Communicating Climate Change*, David C. Holmes and Lucy M. Richardson, eds. (Cheltenham, United Kingdom, Edward Elgar, 2020). أفاد كوك أن Abraham et al. (2014) قد لخص ما جاء في الأدبيات العلمية من أدلة دحضت بقوة الورقات التي تتضمن مزاعم المنكرين، من قبيل مزاعم التبريد في القياسات بواسطة السواتل أو تقديرات انخفاض الحساسية للمناخ. وبالمثل، حاول Benestad et al. (2016) تكرار الاستنتاجات الواردة في ورقات المنكرين ووجد عددا من العيوب مثل الأساليب الإحصائية غير المناسبة، والثنائيات المغلوطة، والاستنتاجات القائمة على المعطيات الفيزيائية الخاطئة.

31 Global Witness, "The climate divide: how Facebook's algorithm amplifies climate disinformation", 28 March 2022.

32 تحليل أجرته إدارة التواصل العالمي، باستخدام بيانات مستقاة من برنامج Talkwalker.

33 Jeffrey A. Hicke and others, "North America", in Intergovernmental Panel on Climate Change, *Climate Change 2022: Impacts, Adaptation and Vulnerability*, Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (Cambridge, United Kingdom, Cambridge University Press, 2022).

34 Mei Li, Gregory Trencher and Jusen Asuka, "*The clean energy claims of BP, Chevron, ExxonMobil and Shell: a mismatch between discourse, actions and investments*", PLOS ONE, issue 17, No. 2 (February 2022).

35 Robert J. Brulle and Carter Werthman, "The role of public relations firms in climate change politics", *Climatic Change*, vol. 169, No. 1–2 (November 2021). وفقا للمؤشر العالمي للمعلومات المضللة، وهو هيئة رقابية غير ربحية، قدم الشركات المعلِنة من قطاع التكنولوجيا 36,7 مليون دولار إلى 98 موقعا إلكترونيا تنشر معلومات مضللة عن المناخ باللغة الإنجليزية في عام 2021. وكشف تقرير صدر في تشرين الثاني/نوفمبر 2022 عن مركز مكافحة الكراهية الرقمية، وهي مجموعة ناشطة في مجال الحملات، أن ما يقرب من نصف مبلغ 23,7 مليون دولار الذي أنفقته شركات النفط والغاز في العامين الماضيين في موقع غوغل وحده على الإعلانات التي تُنشر بجوار نتائج البحث قد استهدف الألفاظ المتعلقة بالاستدامة البيئية. وكشفت التحريات التي أجرتها منظمة InfluenceMap عن 147 25 إعلانا مضللا نشرتها 25 شركة تعمل في قطاع النفط والغاز على منصات فيسبوك في الولايات المتحدة الأمريكية في عام 2020، حيث أنفق ما مجموعه 376 597 9 دولارا. وحتى الآن لم تكن الاستجابة متناسبة مع حجم المشكلة.

36 انظر قرار الجمعية العامة 227/76؛ وقرار مجلس حقوق الإنسان 21/49؛ و European Union External Action, "Tackling disinformation, foreign information manipulation and interference", 27 October 2021.

37 Lucina Di Meco, "Monetizing misogyny: gendered disinformation and the undermining of women's rights and democracy globally", #SheePersisted, February 2023.

38 انظر Andrew Puddephatt, "Social media and elections", Cuadernos de Discusión de Comunicación e Información, No. 14 (Montevideo, UNESCO, 2019); and Julie Posetti and others, "The chilling: global trends in online violence against women journalists", research discussion paper (UNESCO, 2021).

39 EU Disinfo Lab, "The role of "media" in producing and spreading disinformation campaigns", 13 October 2021.

40 انظر United Nations News, "Social media poses 'existential threat' to traditional, trustworthy news: UNESCO", 10 March 2022; and Anya Schiffrin and others, "Finding the funds for journalism to thrive: policy options to support media viability", World Trends in Freedom of Expression and Media Development (Paris, UNESCO, 2022).

41 European Commission, "Shaping Europe's digital future: the 2022 Code of Practice on Disinformation", 4 July 2022.

42 انظر ملاحظات نائبة رئيسة المفوضية الأوروبية المعنية بالقيم والشفافية، فيرا جوروفا، في European Commission, "Code of Practice on Disinformation: new Transparency Centre provides insights and data on online disinformation for the first time", daily news, 9 February 2023. يمكن الاطلاع عليها عبر الرابط التالي: https://ec.europa.eu/commission/presscorner/detail/en/mex_23_723.

43 يمكن الاطلاع عليه عبر الرابط التالي: https://unglobalcompact.org/library/2.

44 Center for Countering Digital Hate and Human Rights Campaign, "Digital hate: social media's role in amplifying dangerous lies about LGBTQ+ people", 10 August 2022.

45 انظر Andrew Puddephatt, "Letting the sun shine in: transparency and accountability in the digital age", World Trends in Freedom of Expression and Media Development (Paris, UNESCO, 2021).

46 Whose Knowledge?, Oxford Internet Institute and The Centre for Internet and Society, State of the Internet's Languages Report (2022).

47 A/HRC/38/35.

48 Billy Perrigo, "Inside Facebook's African sweatshop", *Time*, 17 February 2022.

49 A/HRC/42/50.

50 من الأمثلة البارزة على ذلك Maria Ressa, *How to Stand up to a Dictator* (New York, HarperCollins, 2022); and Max Fischer, *The Chaos Machine* (New York, Little, Brown and Company, 2022).

51 انظر UNICEF, "Young reporters fact-checking COVID-19 information".

52 انظر https://shareverified.com/.

53 تَتَتَبَّع الهيئة غير الربحية المسماة المؤشر العالمي للمعلومات المضللة الإعلانات المنشورة بحذاء المعلومات المضللة. وقد كانت الأمم المتحدة ضحية لهذه الممارسة، حيث وجد المؤشر العالمي للمعلومات المضللة أن إعلانات اليونيسف منشورة بحذاء مقالات مضادة للقاحات، وكذلك إعلانات مفوضية الأمم المتحدة لشؤون اللاجئين التي نشرت بحذاء محتوى مناهض للاجئين.

54 Conscious Advertising Network, manifestos. يمكن الاطلاع عليه عبر الرابط التالي: www.consciousadnetwork.com/the-manifestos/.

55 UNESCO, *Journalism is a Public Good: World Trends in Freedom of Expression and Media Development – Global Report 2021/2022* (Paris, 2022).

56 ITU, "Facts and figures 2021: 2.9 billion people still offline", 29 November 2021. وسيحدد الميثاق الرقمي العالمي الذي ستتباحثه الدول الأعضاء في مؤتمر القمة المعني بالمستقبل، المقرر عقده في عام 2024، المبادئ المشتركة من أجل مستقبل رقمي مفتوح وحر وآمن للجميع (انظر https://www.un.org/techenvoy/ar/global-digital-compact).

57 Krystal Hu, "ChatGPT sets record for fastest-growing user base – analyst note", *Reuters*, 2 February 2023.

58 Center for Countering Digital Hate, "Misinformation on Bard, Google's new AI chat", 5 April 2023; and Tiffany Hsu and Stuart A. Thompson, "Disinformation researchers raise alarms about A.I. chatbots", *The New York Times*, 13 February 2023.

59 يمكن الاطلاع على مشروع المبادئ التوجيهية عبر الرابط التالي: www.unesco.org/ar/internet-conference.