

## التحكم في الإفصاح الإحصائي

### النقاط الرئيسية:

- التحكم في الإفصاح الإحصائي (SDC) هو تقنية تُستخدم لتقييم وتقليل مخاطر إعادة تحديد هوية شخص أو منظمة من خلال تحليل البيانات الدقيقة.
- في القطاع الإنساني، تعتبر البيانات الدقيقة هي البيانات المتعلقة بخصائص السكان التي يتم جمعها من خلال أنشطة مثل المسوحات الأسرية أو تقييم الاحتياجات أو أنشطة الرصد.
- الغرض من تطبيق التحكم في الإفصاح على البيانات الدقيقة الإنسانية هو القدرة على مشاركة البيانات بشكل أوسع بطريقة مسؤولة دون الإضرار بالأشخاص المتضررين.
- يمكن أن تخفض عملية التحكم في الإفصاح الإحصائي من مخاطر إعادة التحديد إلى مستوى مقبول، ولكن قد يختلف عبء المخاطر بناءً على السياق الذي تحدث فيه الاستجابة الإنسانية.
- تشمل القيود الشائعة المتعلقة بمشاركة البيانات مع الجهات المانحة: (i) نقص الأطر التنظيمية لإدارة البيانات الحساسة غير الشخصية بمسؤولية؛ (ii) فجوات في القدرات؛ و (iii) تحديد الغرض.
- لبدء استخدام التحكم في الإفصاح الإحصائي، يجب على المنظمات الاستثمار في (١) العثور على الأداة المناسبة، (٢) إعداد سير العمل، و (٣) تحسين الممارسة بمرور الوقت من خلال التعلم المستمر.

### ما هي البيانات الدقيقة الإنسانية؟

البيانات المتعلقة بخصائص وحدات السكان (مثل الأفراد، الأسر أو المنشآت) التي يتم جمعها من خلال تعداد، مسح أو تجربة تُعرف في الإحصائيات باسم «البيانات الدقيقة»<sup>١</sup>. في الاستجابة الإنسانية، يتم جمع هذا النوع من البيانات من خلال أنشطة مثل تقييم الاحتياجات متعددة القطاعات (MSNA)، المسوحات الأسرية، وأنشطة تقييم الاحتياجات أو الرصد الأخرى. تشكل هذه البيانات جزءاً متزايد الأهمية من حجم البيانات في القطاع الإنساني، وتصبح أكثر حيوية لتحديد احتياجات وآراء الأشخاص المتأثرين بالأزمات<sup>٢</sup>. لذلك، من الضروري أن تفهم المنظمات الإنسانية كيفية تقييم وإدارة حساسية هذه البيانات لضمان استخدامها الكامل وتأثيرها في سياقات الاستجابة المختلفة.

في شكلها الخام، قد تحتوي البيانات الدقيقة على كل من البيانات الشخصية والبيانات غير الشخصية حول مجموعة متنوعة من المواضيع، بما في ذلك مواضيع حساسة مثل التعرض للعنف القائم على النوع الاجتماعي، الأمراض المعدية، وغيرها من القضايا التي قد يتم تسجيلها في حقول النصوص الحرة. تعترف معظم المنظمات الإنسانية بحساسية البيانات الشخصية مثل الأسماء، البيانات البيومترية، أو أرقام الهوية، وتقوم بإخفاء هوية مجموعات البيانات وفقاً لذلك كمسألة ممارسة معيارية. ومع ذلك، لا يزال من الممكن في كثير من الأحيان إعادة تحديد هوية المجهين الأفراد أو المنظمات من خلال دمج الإجابات على أسئلة مختلفة، حتى بعد تطبيق هذا «إخفاء الهوية».

<sup>١</sup> لمزيد من المعلومات حول مخاطر الإفصاح والاعتبارات التقنية المتعلقة بتقييم وإدارة هذه المخاطر من خلال التحكم الإحصائي في الإفصاح (SDC)، يمكنكم الاطلاع على دليل الممارسة: التحكم الإحصائي في الإفصاح للبيانات الجزئية، المتاح هنا: <https://sdcppractice.readthedocs.io/en/latest/index.html>.

<sup>٢</sup> في وقت كتابة هذا النص، أظهر البحث عن كلمة «survey» على منصة Humanitarian Data Exchange عدد ١١٩٨ نتيجة من بين ٩٨٠٥ مجموعات بيانات موجودة على المنصة؛ بينما أظهر البحث عن كلمة «assessment» عدد ١٣٩٩ نتيجة.

## إعادة التحديد وخطر الإفشاء

يمكن أن تسمح سلسلة من نقاط البيانات بإعادة تحديد الهوية، سواء كانت بشكل منفرد أو عند دمجها مع فهم سياقي أساسي. كما يمكن للتقنيات المتقدمة لتحليل البيانات استخراج رؤى أكثر حساسية مما قد يكون مرئيًا من خلال التحليل الأساسي، مما يزيد من الحساسية المحتملة للبيانات الدقيقة في القطاع الإنساني.

هناك ثلاثة أشكال معترف بها بشكل شائع لخطر الإفشاء، كل منها يمكن أن يظهر في البيانات الدقيقة الإنسانية:

- **إفشاء الهوية:** عندما يمكن ربط فرد معروف بسجل بيانات منشور
- **إفشاء السمّة:** عندما يمكن تحديد خاصية جديدة للفرد بناءً على المعلومات المتاحة في البيانات المنشورة
- **الإفشاء الاستنتاجي:** عندما يمكن تحديد خاصية للفرد بدقة أكبر باستخدام البيانات المنشورة مما كان ممكنًا في السابق

## التحكم في الإفشاء الإحصائي

إحدى الطرق لتقييم وتقليل خطر إعادة التحديد في البيانات الدقيقة هي التحكم في الإفشاء الإحصائي (SDC). SDC هي تقنية تُستخدم في الإحصاء لتقييم وتقليل خطر إعادة تحديد الشخص أو المنظمة من نتائج تحليل بيانات المسح أو البيانات الإدارية، أو في نشر البيانات الدقيقة. تم استخدام هذه التقنية بشكل أساسي من قبل المكاتب الإحصائية الوطنية (NSOs) والمنظمات الإحصائية الأخرى فيما يتعلق ببيانات التعداد.

من خلال تطبيق SDC، ستأثر دائمًا القيمة المعلوماتية الإجمالية لمجموعة البيانات، ويُعتبر تحقيق توازن مناسب بين خطر إعادة التحديد وفقدان قيمة المعلومات أمرًا حاسمًا في أي تطبيق لـ SDC. عند تحديد التوازن المناسب بين المخاطر والمنفعة، من الضروري مراعاة الاستخدامات المختلفة المحتملة لمجموعة البيانات والسياق الذي تم فيه جمع البيانات.

عادةً ما يتكون إجراء SDC من ثلاث خطوات:

### ١. تقييم المخاطر

في الخطوة الأولى، قم بتقييم احتمال حدوث إفشاء ناجح للمستجيبين الفرديين ضمن مجموعة بيانات معينة. يعتمد قبول نسبة المخاطر لمجموعة البيانات على السياق المتعلق بالبيانات. على سبيل المثال، في بيئة الصراع، تكون نسبة المخاطر المسموح بها عادةً أقل مما تكون عليه في استجابة لكارثة طبيعية.

### ٢. تطبيق طرق التحكم في الإفشاء الإحصائي

في الخطوة الثانية، يتم تمرير مجموعة البيانات من خلال عملية SDC الفعلية، التي تقلل من خطر إعادة التحديد عن طريق تطبيق واحدة أو أكثر من طرق إخفاء الهوية. تنقسم هذه الطرق إلى فئتين: الطرق التعديلية، التي لا تقمع القيم في مجموعة البيانات ولكن تقوم بتعديل القيم للحد من خطر الإفشاء من خلال خلق عدم يقين حول القيم الحقيقية، أو الطرق غير التعديلية التي تقلل التفاصيل في البيانات من خلال التعميم أو قمع بعض القيم (أي إخفاء) دون تشويه بنية البيانات.

### ٣. إعادة تقييم المخاطر وقياس فقدان المعلومات

الخطوة الأخيرة في العملية هي قياس فقدان المعلومات الناتج عن تطبيق طرق SDC المختلفة على مجموعة البيانات. يتم مقارنة القيمة المعلوماتية لمجموعة البيانات الأصلية مع القيمة المعلوماتية النهائية، ويتضمن ذلك أيضًا إعادة تقييم خطر إعادة التحديد للتأكد من أنه تم تخفيضه بشكل كافٍ.

الغرض من تطبيق التحكم في الإفشاء الإحصائي (SDC) على البيانات الدقيقة في القطاع الإنساني هو تمكين نشر بيانات المسح وتقييم الاحتياجات على نطاق أوسع بطريقة مسؤولة. من خلال استخدام SDC كوسيلة لتقييم وتقليل حساسية البيانات، يمكن للمنظمات الإنسانية نشر بيانات المسح وتقييم الاحتياجات بشكل مسؤول للمساهمة في توجيه جهود الاستجابة الشاملة.

<sup>٢</sup> مزيد من المعلومات حول مخاطر الإفشاء والاعتبارات التقنية المتعلقة بتقييم وإدارة هذه المخاطر من خلال SDC، راجع دليل الممارسة لإدارة التحكم في الكشف الإحصائي للبيانات الدقيقة، والمتاح هنا: <https://sdcppractice.readthedocs.io/en/latest/index.html> (<https://sdcppractice.readthedocs.io/en/latest/index.html>)

## تطبيق التحكم في الإفشاء الإحصائي (SDC) على البيانات المشتركة على منصة HDX

منذ بداية عام ٢٠١٨، قام مركز البيانات الإنسانية بإجراء تقييم للمخاطر لـ ٥٩ مجموعة بيانات تم تحميلها على منصة تبادل البيانات الإنسانية (HDX). واعتُبر أن خطر إفشاء هويات المستجيبين في ٣٨ من تلك المجموعات كان مرتفعاً جداً للنشر على HDX. وافق المساهمون في ١٤ من هذه المجموعات على تطبيق SDC على بياناتهم لخفض مستوى المخاطر. قام فريق HDX بتطبيق SDC على هذه المجموعات الـ ١٤، مما أدى إلى خفض مستوى المخاطر إلى مستوى مقبول (أي ٥٪ أو أقل). بالنسبة لـ ٥ من هذه المجموعات الـ ١٤، يعني ذلك أنه يمكن نشرها بشكل علني بعد إخفاء الهوية. أما المجموعات التسع المتبقية، فقد تم إما إزالتها أو مشاركتها بشكل خاص على HDX، وكذلك الحال بالنسبة لـ ٢٤ مجموعة بيانات التي كانت عالية المخاطر والتي لم يوافق المساهمون على إجراء SDC عليها. بالنسبة لهذه المجموعات الـ ٢٤، اتخذت العديد من المنظمات تدابير خاصة بها لخفض خطر إعادة التعرف على الهوية، والتي تضمنت أحياناً إزالة المتغيرات الحساسة غير الضرورية بالكامل.

## في إدارة البيانات الإنسانية (SDC) تطبيقات التحكم في الإفشاء الإحصائي

في أوائل عام ٢٠١٩، أجرى مركز البيانات الإنسانية مقابلات مع سبع منظمات إنسانية تقوم بإجراء استطلاعات وتقييمات احتياجات بشكل منتظم لفهم الممارسات الحالية لإدارة البيانات الجزئية. بينما تمتلك بعض المنظمات مثل المفوضية السامية للأمم المتحدة لشؤون اللاجئين (UNHCR) (انظر دراسة الحالة أدناه) نهجاً متقدماً نسبياً وخبرة كبيرة داخل المنظمة في إجراء SDC على أشكال مختلفة من البيانات الجزئية، فإن معظم المنظمات التي تم مقابلتها تحتاج إلى دعم للقيام بهذا العمل.

### التنظيم المسؤول وإدارة البيانات الجزئية عن اللاجئين تجربة المفوضية السامية للأمم المتحدة لشؤون اللاجئين (UNHCR)

تقوم المفوضية السامية للأمم المتحدة لشؤون اللاجئين (UNHCR) بجمع البيانات بانتظام عن اللاجئين والسكان الآخرين الذين يقعون تحت ولايتها. تُستخدم هذه البيانات لتقييم الاحتياجات ونقاط الضعف، وتوجيه البرامج، وتحسين استهداف المساعدات. وعلى الرغم من أن هذه البيانات لم تُحفظ تقليدياً في صيغ ومواقع تجعلها قابلة للاسترجاع بسهولة للاستخدام في المستقبل، فإن UNHCR تعمل الآن على إنشاء مكتبة بيانات جزئية داخلية وخارجية. من خلال إنشاء هذه المستودعات عبر الإنترنت التي ستتيح الوصول العام إلى البيانات الجزئية للمستخدمين الداخليين والخارجيين، تهدف UNHCR إلى تمكين استخدام أكثر شمولاً للبيانات من قِبل مجموعة متنوعة من أصحاب المصلحة ومنع التكرار في جهود جمع البيانات في المستقبل.

بينما تحمل الإتاحة العامة للبيانات الجزئية العديد من الفوائد المحتملة، فإنها تأتي أيضاً مع مخاطر محتملة. يمكن أن يؤدي النشر دون اتخاذ تدابير التحكم في الإفشاء المناسبة إلى تمكين المتسللين من تحديد الكيانات (الأفراد أو الأسر) التي تتم مشاركة بياناتها، حتى إذا تم إزالة المعلومات المباشرة مثل الأسماء والعناوين. وفقاً لسياسة حماية البيانات الخاصة بـ UNHCR، يجب حماية هوية الأشخاص المعنيين، وبالتالي يجب أن تكون مجموعات البيانات مجهولة الهوية بشكل صحيح قبل أن يتم مشاركتها. بيانات UNHCR حساسة بشكل خاص، لأنها تتعلق بفئات من الناس الضعفاء بشكل خاص، وحمايتهم من الأهمية القصوى.

لضمان الحماية والنشر المسؤول للبيانات الجزئية، تستخدم UNHCR تطبيق sdcMicro في R لحساب مخاطر إعادة تحديد الهوية لهذه البيانات قبل نشرها. يتم إدارة العملية من قبل فريق تنظيم البيانات في UNHCR، والذي يعمل بالتعاون مع أصحاب البيانات لتحديد المتغيرات الرئيسية، وتقييم حساسية البيانات، وتحديد مستوى الخطر المقبول لمجموعة البيانات المعنية. بعد عملية إخفاء الهوية، يتم مقارنة البيانات المعدلة بالبيانات الأصلية وتقييم فقدان المعلومات. إذا رأى صاحب البيانات أن بعض المتغيرات المعدلة ضرورية لمستهلقي البيانات، يمكن تعديل أساليب التحكم في الإفشاء وفقاً لذلك. على سبيل المثال، في حالة المسوحات الغذائية الموسعة الموحدة (SENS)، قرر فريق التنظيم عدم تطبيق التجميع في فئات عمرية كانت تُطبق عادةً لأن هذه الفئات كانت أساسية لتوصيف سوء التغذية حسب العمر بالسنوات والأشهر للأطفال. احتفظ الفريق بمتغير العمر لكنه استبعد تاريخ الميلاد وتاريخ المسح. أدى ذلك إلى سيناريو خطر مقبول مع الحفاظ على فائدة البيانات لأخصائيي التغذية.

تواصل UNHCR الاستثمار في هذه العملية من خلال تعزيز فريق التنظيم وزيادة الخبرة التقنية في تقنيات إخفاء الهوية داخل المنظمة. وفقاً للخطة الحالية، ستكون مكتبة البيانات الجزئية التابعة لـ UNHCR جاهزة للعمل ومجهزة ببيانات النزوح القسري بحلول نهاية عام ٢٠١٩.

من خلال العمل مع المساهمين في البيانات مثل REACH (انظر دراسة الحالة أدناه) لتطوير خدمة SDC موثوقة وأمنة، يهدف مركز البيانات الإنسانية إلى دعم المشاركة المسؤولة لهذه البيانات وإثبات قيمة تقنيات أكثر قوة لتقييم المخاطر وإخفاء الهوية. يساعد التعرض لهذه التقنيات المنظمات الإنسانية المختلفة على تحديد الأدوات والأساليب التي يمكنها دمجها في عمليات إدارة البيانات الخاصة بها، بينما يساهم أيضًا في توسيع المعرفة داخل القطاع حول كيفية إدارة ومشاركة البيانات الجزئية بشكل أكثر مسؤولية في السياقات الإنسانية.

## الفرص والتحديات لإدماج SDC في سير العمل الخاص بالمنظمة تجربة REACH

بدأت REACH في استكشاف إمكانيات SDC في يونيو ٢٠١٨، عندما طبق فريق HDX لأول مرة حزمة sdcMicro R على مجموعة بيانات قامت REACH بتحميلها على المنصة. تشمل أنواع البيانات التي تم تطبيق SDC عليها من قبل فريق HDX لـ REACH استطلاعات الأسر ومقابلات المخبرين الرئيسيين (والبيانات الوصفية المرتبطة بها). لم تقم REACH بعد بتطبيق SDC مباشرة ولكنها تدرس المتطلبات اللازمة للقيام بذلك.

استنادًا إلى الخبرة حتى الآن، تقترح REACH على المنظمات المهتمة بإدماج SDC في سير عملها النظر في الأسئلة التالية:

- هل هذه هي المنهجية المناسبة لعمليات إدارة البيانات الجزئية الحالية لديك؟
- إلى أي مدى يؤدي تطبيق SDC إلى خفض صحة وفائدة البيانات؟
- كيف يؤثر تطبيق SDC على الشفافية؟
- كيف يمكن ضمان أن لا يعتمد الموظفون بشكل مفرط على نتائج تقييم مخاطر الكشف عن SDC، وضمان استمرارهم في التفكير النقدي بشأن المخاطر المحتملة لأنواع البيانات المختلفة؟

قد حددت REACH أنه سيكون من الممكن تقنيًا تنفيذ الجوانب الفنية لـ SDC بسهولة نسبية سواء في المقر الرئيسي أو على مستوى الميدان. في المقر الرئيسي، سيكون هذا يعني تشغيل برنامج نصي على جميع مجموعات البيانات التي تنتجها أو تنشرها REACH. وعلى مستوى الميدان، سيعني ذلك حث الفرق البلدان على استخدام sdcMicro أو أداة مماثلة على جميع مجموعات البيانات المنتجة في البلد.

إلى جانب الجوانب الفنية لـ SDC، ترى REACH أن التحدي المحتمل أو نقطة الاختناق تكمن في الجزء اليديوي من العملية، حيث يجب على الموظفين تحديد ما إذا كانت تقنية التحكم في الكشف مناسبة، وأي المتغيرات يجب إزالتها أو إخفاؤها، وكيفية تفسير نتائج العملية والتواصل بشأنها. هذه القرارات تستغرق وقتًا وتتطلب فهمًا للسياق الذي ترتبط به البيانات.

على المدى القريب، ستواصل REACH التعاون مع فريق HDX لإجراء SDC على بيانات الاستطلاعات والتقييمات قبل نشرها على HDX. سنتيح هذه التجربة لـ REACH تحديد أفضل السبل لإدماج SDC في سير العمل الخاص بها على المستوى العالمي والمستوى الوطني في المستقبل.

## مجالات الاستثمار: زيادة استخدام SDC في إدارة البيانات الجزئية الإنسانية

الاستثمار في القدرة والبنية التحتية اللازمة لتطبيق SDC يمكن المنظمات من تحديد المخاطر المرتبطة بمشاركة مجموعات بيانات الاستطلاعات وتقييم الاحتياجات. يوصي مركز البيانات الإنسانية بالاستثمار في المجالات الثلاثة التالية لاعتماد SDC بنجاح من قبل المنظمات الإنسانية.

### ١. الأدوات

قامت العديد من المنظمات بتطوير أدوات مفتوحة المصدر لإجراء SDC وجعلتها متاحة مجانًا عبر الإنترنت. يستخدم مركز البيانات الإنسانية ومنظمات إنسانية أخرى استشارها المركز أثناء بحثه حاليًا **sdcMicro**. اختار المركز **sdcMicro** نظرًا لقابليته للتوسع ولأنه مجاني ومفتوح المصدر. تشمل الأدوات المجانية والمفتوحة المصدر الأخرى **μArgus** و **ARX**. عند اختيار الأداة المناسبة لإجراء SDC، يجب على المنظمات مراعاة قدرة الأداة على التعامل مع بيانات كبيرة، والموازنة بين المخاطر والفائدة، وسهولة استخدام واجهة المستخدم للأداة.

### ٢. سير العمل

إلى جانب اختيار الأداة المناسبة لـ SDC، يعد إعداد سير عمل واضح لتطبيق الأداة المختارة أمرًا أساسيًا. يتطلب SDC مشاركة أدوار مختلفة في العملية، بما في ذلك متخصص تقني لتطبيق الأداة، وشخص لديه فهم للسياق الذي ترتبط به البيانات لتحديد مستوى المخاطر المقبول، وآخرين. سيساعد سير العمل المنظم بشكل جيد في تحسين كفاءة العملية ومنع سوء تفسير أو الاعتماد المفرط على نتائج SDC.

### ٣. التعلم المستمر

مع تطبيق المنظمات لـ SDC، ستتعلم بمرور الوقت حول حساسية المتغيرات الرئيسية المختلفة، ومستوى المخاطر المناسب الذي يجب السعي لتحقيقه، ومستوى فقدان المعلومات المقبول، ومختلف الاعتبارات الأخرى التي يجب موازنتها في العملية. يعد الاحتفاظ بسجل لكل تطبيق لـ SDC وتوثيق الدروس المستفادة أمرًا مهمًا، خاصة في الفترة الأولى من إدخال SDC. يعتبر مشاركة هذه الأفكار داخليًا عبر الفرق، وعند الاقتضاء، مع المجتمع الإنساني الأوسع، طريقة رائعة لدعم إدارة أكثر اتساقًا ومسؤولية للبيانات الجزئية في القطاع.

لتتعرف على المزيد حول عمل المركز على SDC للبيانات الجزئية الإنسانية، يمكن الاتصال بـ [centrehumdata@un.org](mailto:centrehumdata@un.org).

المتعاونون: المفوضية السامية للأمم المتحدة لشؤون اللاجئين (UNHCR)، مبادرة REACH

ينشر مركز البيانات الإنسانية «(المركز)» بالتعاون مع الشركاء الرئيسيين، سلسلة مذكرات إرشادية بشأن مسؤولية البيانات في العمل الإنساني خلال عامي ٢٠١٩ و ٢٠٢٠. وتتبع سلسلة المذكرات الإرشادية نشر إرشادات مسؤولية البيانات لمكتب تنسيق الشؤون الإنسانية في مارس ٢٠١٩. من خلال هذه السلسلة، يهدف المركز إلى تقديم إرشادات إضافية حول القضايا والعمليات والأدوات المحددة لمسؤولية البيانات في الممارسة العملية. وقد تم إعداد هذه السلسلة بفضل الدعم السخي من المديرية العامة للحماية المدنية وعمليات المساعدة الإنسانية الأوروبية (ECHO DG).

يغطي هذا المستند أنشطة المساعدات الإنسانية المنفذة بمساعدة مالية من الاتحاد الأوروبي. لا ينبغي بأي شكل من الأشكال اعتبار الآراء الواردة هنا تعبيرًا عن الرأي الرسمي للاتحاد الأوروبي، كما أن المفوضية الأوروبية ليست مسؤولة عن أي استخدام قد يتم للمعلومات التي يحتويها.



هذا المشروع ممول جزئيًا من الاتحاد الأوروبي