



الشبكة الدولية لاستقصاءات الأسر المعيشية

بثّ ونشر ملفات البيانات الجزئية مبادئ وأنظمة وممارسات

أوليفيه دوبريز وإيرني بويكو

وثيقة العمل رقم 005 للشبكة الدولية لاستقصاءات الأسر المعيشية IHSN
أغسطس / آب 2010

بثّ ونشر ملفات البيانات الجزئية مبادئ وأنظمة وممارسات

أوليفيه دوبريز وإيرني بويكو

وثيقة العمل رقم 005 للشبكة الدولية لاستقصاءات الأسر المعيشية IHSN

أغسطس / آب 2010

خلاصة

يواجه منتجو البيانات من جميع البلدان طلباً متنامياً على البيانات الجزئية. وتمثل معرفة الطريقة الأفضل لبث هذه البيانات تحدياً كبيراً، ويكتسي هذا التحدي طابعاً تقنياً، إذ تتطلب العملية وضع إجراءات توثيق البيانات، وفهرستها وبثها أو نشرها، كما أنه يكتسي أيضاً طبيعة قانونية وأخلاقية. وإذا كان منتجو البيانات على دراية تامة بقوة البيانات الجزئية وأهميتها، فإن عليهم أن يوازنوا بين الطلب عليها وضرورة تأمين سرية المعلومات التي يقدمها الأشخاص المستجيبون. وترعى هذا المطلب التشريعات الوطنية المتعلقة بالإحصاءات وسرية المعلومات، وغالباً ما يشكل التزاماً تجاه المستجيبين أثناء جمع المعلومات. ويستدعي ذلك وضع سياسات وإجراءات تحدد شروط النفاذ إلى البيانات الجزئية. ويتضمن هذا المستند على وصف عام لهذه السياسات والإجراءات، كما يحصي الممارسات السليمة في هذا الشأن.

المؤلفان

عمل إيرني بويكو Ernie Boyko لفترة طويلة في إحصاءات كندا، حيث أدار تبعاً أقسام الإحصاءات الزراعية، والتخطيط المتكامل وأنظمة التدبير، والبث المعلوماتي للبيانات، ثم ترأس عمليات تعداد عام 1991. كما أشرف على أعمال مبادرة إتاحة البيانات IDD. ويشغل إيرني بويكو منصب عضو فاعل في المؤسسة الكندية لمستخدمي البيانات العامة AUCDP، والمؤسسة الدولية لخدمات المعلوماتية وتقنياتها في العلوم الاجتماعية IASSIST.

أوليفيه دوبريز Olivier Dupriez، خبير معتمد في الاقتصاد والإحصاء في مجموعة إدارة البيانات المتعلقة بالتنمية في البنك الدولي. كما يشغل منصب منسق في الشبكة الدولية لاستقصاءات الأسر المعيشية IHSN. ويضطلع إداً بمهام تنسيق برامج المساعدة التقنية لعدد من البلدان في المجالات المتعلقة بتوثيق البيانات الجزئية وبثها.

كلمة شكر

أعدّ هذا المستند من طرف شبكة IHSN بدعم مالي من آلية DGF (مرفق المنح الإنمائية) من البنك الدولي، وقد حولت أمانة PARIS21 في منظمة التعاون والتنمية في الميدان الاقتصادي OECD المنحة رقم 4001009-06 لهذه الغاية.

أعدّه كل من إيرني بويكو وأوليفيه دوبريز بمساعدة أشخاص كانوا مصدرراً للمعلومات أو قدّموا ملاحظات واقتراحات بشأن المستند، وهم: فرانسوا فونتونو -teneau-François Fon (PARIS21، منظمة التعاون والتنمية في الميدان الاقتصادي)، وجوليا لاين Julia Lane (مركز الرأي الوطني للبحوث Research Center National Opinion، جامعة شيكاغو)، وجوهان ميستيان Johan Mistiaen (البنك الدولي)، ودنيس تريوين Dennis Trewin وويندي واتكينس Wendy Watkins (جامعة كارلتون، كندا).

يضمّ هذا المستند أيضاً مناقشات مع عدد كبير من الزملاء في الوكالات المنتسبة إلى شبكة IHSN، ومع القائمين على الإحصاءات الرسمية في عدد من الدول. أعدّه للنشر جون رايت John Wright، وصمّمه روميل ريكو Rhommell Rico.

يُسمح بتداول وثيقة العمل هذه وباستخدامها. ولكن يحظر كل استخدام تجاري لنسخة عنها في هذا الإطار.

إن هذه الوثيقة (أو أي نسخة حديثة عنها) متاحة على موقع IHSN الإلكتروني: www.ihsn.org.

المراجع

دوبريز، أوليفيه وإيرني بويكو. 2010. بث ملفات البيانات الجزئية. تعريف بالسياسات والإجراءات، الشبكة الدولية لاستقصاءات الأسر المعيشية (IHSN)، وثيقة العمل رقم 005.

إن النظريات، والتحليلات والآراء الموجودة في هذا المستند لا تعبّر إلا عن صاحبها/ أصحابها، ولا تعكس بالضرورة رأي الوكالات المنتسبة إلى شبكة IHSN أو أمانتها.

فهرس المحتويات

ii	خلاصة
ii	المؤلفان
iii	كلمة شكر
iv	فهرس المحتويات
viii	مقدمة
1	1. ما هي البيانات الجزئية؟
2	1.1 مفهوم البيانات الجزئية
2	1.2 شكل تخزين ملفات البيانات الجزئية وبثها
5	1.3 أي نسخة ينبغي بثها من ملفات البيانات؟
5	1.4 ما هي العناصر الحساسة في محتوى البيانات الجزئية؟
5	1.5 ما هي أنواع ملفات البيانات الجزئية الرئيسية التي تبث؟
7	1.6 هل ثمة بدائل لمشاركة البيانات الجزئية؟
9	2. ما هي البيانات الوصفية؟
9	ما هي البيانات الوصفية ذات الجودة؟
10	المعايير والممارسات السليمة المتعلقة بالبيانات الوصفية
14	3. ما هي الحجج التي تدعم بث البيانات الجزئية؟
14	3.1 دعم البحث
14	3.2 تعزيز مصداقية الإحصاءات الرسمية
14	3.3 تحسين موثوقية البيانات وثباتها
14	3.4 تقليص تكرار البيانات
14	3.5 تحسين مردود الاستثمار
15	3.6 جمع الأموال للدراسات الإحصائية
15	3.7 تقليص تكلفة نشر البيانات
15	3.8 احترام الإلتزامات القانونية والتعاقدية
15	3.9 تشجيع تطوير أدوات جديدة لاستخدام البيانات
17	4. ما هي التكاليف والمخاطر المرتبطة ببث البيانات الجزئية وكيف يمكن ضبطها؟
17	4.1 المسائل الأخلاقية والحفاظ على سرية المستجيبين
17	4.2 النواحي القانونية
19	4.3 التعرض إلى الانتقاد والمعارضة
19	4.4 التكاليف
20	4.5 خسارة الحصرية
20	4.6 القدرات التقنية
21	5. إلى من تتوجه هذه البيانات الجزئية؟
23	6. ما هي الشروط التي تحكم بث البيانات؟
24	6.1 الأساس التشريعي
24	6.2 الشروط المطبقة على ملفات البيانات الجزئية المتاحة للاستخدام العام
24	6.3 الشروط التي تطبق على الملفات الخاضعة لترخيص
25	6.4 شروط خاصة بمراكز البيانات المحمية
25	6.5 إدارة مخالفات الباحثين
29	7. ماذا نعني بـ“إغفال” البيانات الجزئية؟
29	7.1 مفاهيم مرتبطة بضبط الكشف عن الإحصاءات
29	7.2 سيناريوهات الكشف
30	7.3 تقييم خطر الكشف
30	7.4 تقنيات ضبط الكشف عن السرية الخاصة بملفات البيانات الجزئية
32	7.5 إيجاد الحل الوسط بين خطر الكشف وخسارة المعلومات

- 7.6. توثيق عملية ضبط الكشف عن البيانات الإحصائية 32
8. هل ينبغي أن يكون النفاذ إلى البيانات الجزئية مجانياً أم مدفوع الأجر؟ 34
- 8.1. مثال عن بلدين 34
- 8.2. النفاذ مدفوع الأجر أو مجاني؟ 34
9. في أي وقت من دورة البث ينبغي أن تجعل البيانات الجزئية عامة؟ 36
10. ما هي المتطلبات اللازمة من حيث البنية التحتية التقنية؟ 37
11. ما هي المتطلبات المؤسسية المتعلقة بإتاحة ملفات البيانات الجزئية؟ 40
12. ما السبيل إلى التشجيع على استخدام ملفات البيانات الجزئية؟ 42
- المراجع 50
- المواقع الإلكترونية 52

الملحقات

- الملحق 1: طلب نفاذ إلى مجموع البيانات الخاضعة لترخيص في إطار مشروع بحث محدد 43
- الملحق 2: نموذج سياسة النفاذ إلى مركز بيانات محمي 45
- الملحق 3: طلب نفاذ إلى مجموع البيانات الخاضعة لترخيص في إطار مشروع بحث محدد 47

فهرس الرسوم

- الرسم 1 - مقتطف من ملف بيانات بصيغة ASCII ثابتة 2
- الرسم 2 - مقتطف من ملف بيانات بصيغة Stata 4
- الرسم 3 - دورة حياة استقصاء 5

فهرس الإطارات

- الإطار 1 استقصاء واحد، ومنتجات عديدة 7
- الإطار 2 دراسة الدخل في لوكسمبورغ - Luxembourg Income Study - LISSY 8
- الإطار 3 إلى من يتوجه معيار DDI؟ 11
- الإطار 4 إلى من يتوجه Dublin Core؟ 12
- الإطار 5 لغة XML 13
- الإطار 6 الالتزام القانوني بث البيانات الجزئية؛ مثال المركز الوطني للإحصاءات الصحية (الولايات المتحدة) 15
- الإطار 7 الترويج لتطوير التطبيقات المركبة mashups عن طريق بث البيانات المفتوحة وواجهات برمجة التطبيقات 16
- الإطار 8 أمثلة على التشريعات في مسألة السرية 19
- الإطار 9 مثال عن التعهد بالسرية 22
- الإطار 10 شروط النفاذ والاستخدام المطبقة على ملفات البيانات الجزئية المتاحة للاستخدام العام 25
- الإطار 11 كيف يتم ذكر ملف بيانات إلكتروني 26
- الإطار 12 شروط النفاذ إلى ملفات البيانات الخاضعة لترخيص واستخدامها 27
- الإطار 13 العقد الإطاري 28
- الإطار 14 قائمة النقاط التي ينبغي التحقق منها في إطار تقييم مختلف الحالات ومختلف مخاطر الكشف 30
- الإطار 15 الإجراءات المطبقة على عينات البيانات الجزئية المتاحة للاستخدام العام لإحصاء عام 2000 33
- الإطار 16 سياسة آجال نشر البيانات في المركز الوطني للإحصاءات الصحية (الولايات المتحدة) NCHS 36
- الإطار 17 مجموعة أدوات إدارة البيانات الجزئية (الشبكة الدولية لاستقصاءات الأسر المعيشية IHSN) 37

الاختصارات

المكتب الأسترالي للإحصاءات	ABS
جرد الجماعة الأمريكية: مسح سنوي في الولايات المتحدة يستجوب عينة من السكان	ACS
واجهة برمجة التطبيقات	API
Application Programming Interface	ASCII
American Standard Code for Information Interchange	CDR
الترميز الأمريكي المعياري لتبادل المعلومات	SDC
مركز بيانات البحث (إحصائيات كندا)	CEE-ONU
مراقبة الكشف عن الإحصائيات	CENEX
اللجنة الاقتصادية للأمم المتحدة في أوروبا	CESSDA
مشروع أوروبي يهدف إلى وضع قائمة بالممارسات السائدة في أوروبا في ما يتعلق بمسألة حماية البيانات	CSO
Disclosure Control	CURF
Centre of Excellence for Statistical	DCMI
المجلس الأوروبي لأرشيف بيانات العلوم الاجتماعية	DDI
Council of European Social Science Data Archives	DHS
المكتب المركزي للإحصاء في أيرلندا	DSNU
Central Statistics Office	EU
ملفات تسجيل فردية سرية	FMGD
Confidentialised Unit Record Files	GPS
مبادرة Dublin Core للبيانات الوصفية	HTML
Dublin Core Metadata Initiative : منظمة تدعم الأنشطة المتعلقة بإعداد رسم للبيانات الوصفية العامة	HTTP
Documentation Initiative	ICPSR
تحالف مبادرة توثيق البيانات - مشروع لوضع معايير التوثيق التقنية في العلوم الاجتماعية	IDD
Data	IHSN
الاستقصاءات الديموغرافية والصحية	ISO/IEC
Demographic and Health Surveys : برنامج عالمي للأبحاث الديموغرافية والصحية	JSI
الشعبة الإحصائية للأمم المتحدة	LIS
الولايات المتحدة الأمريكية	MCRDC
ملف البيانات الجزئية العمومي	MICS
نظام تحديد المواقع	MIT
Global Positioning System : نظام تحديد المواقع يعمل على الصعيد العالمي	NCHS
لغة ترميز النصوص التشعبية	NCSA
HyperText Markup Language : شكل المعطيات مصمم لتمثيل صفحات شبكة الأنترنت	NDE
بروتوكول نقل النصوص التشعبية	NORC
Hypertext Transfer Protocol	NSD
اتحاد أبحاث العلوم السياسية والاجتماعية بين الجامعات	UN
Consortium for Political and Social Research	OAIS
Inter-University	OCLC
مبادرة إتاحة البيانات (إحصاءات كندا)	Online
Initiative de démocratisation des données	OECD
الشبكة الدولية لاستقصاءات الأسر المعيشية	ONG
International Household Survey Network	ONS
المنظمة العالمية في تطبيق المعايير / اللجنة الكهروتقنية الدولية	PDF
واجهة لنشر الوظائف وعرضها	PUMA
Job Submission Interface	PUMS
قاعدة بيانات ناتجة عن دراسة المداخيل	SAS
Luxemburg Income Study	SSN
مركز ميشيغان للبحوث وبيانات السكان	
Michigan Census Research Data Center	
منهجية الاستقصاء المتعدد المؤشرات	
Multiple Indicator Cluster Surveys	
معهد ماسشوستس التقني	
Institut de technologie du Massachusetts	
المركز الوطني للإحصاءات الصحية (الولايات المتحدة)	
National Center for Health Statistics	
المركز الأمريكي لتشغيل التطبيقات عالية الأداء	
National Center for Supercomputing Applications	
مركز البيانات المحمية	
NDE	
المركز الوطني لبحوث الرأي (جامعة شيكاغو)	
National Opinion Research Center	
أرشيف بيانات الزواج في العلوم الاجتماعية	
Norwegian Social Science Data Services	
الأمم المتحدة	
United Nations	
نموذج مرجعي لنظام أرشفة معلومات مفتوح	
Open Archival Information System	
منظمة بحثية عالمية غير ربحية تساعد المكتبات للوصول إلى المعلومات في جميع أنحاء العالم	
Computer Library Center	
منظمة التعاون والتنمية في الميدان الاقتصادي	
OECD	
منظمة غير حكومية	
ONG	
المكتب الوطني للإحصاء	
ONS	
هيئة الوثائق المحمولة - تم تطويرها من قبل أدوبي	
Portable Document Format	
وحدات البيانات الجزئية المعجمة	
PUMA	
عينات من البيانات الجزئية المعجمة	
PUMS	
برمجية	
Statistical Analysis System	
نظام إحصائي وطني	
SSN	

Statistics New Zealand	معهد الإحصاءات النيوزيلندي	SNZ
XML Simple Object Access Protocol	بروتوكول نفاذ قائم على لغة	SOAP
Structured Query Language	لغة الاستعلامات البنائية	SQL
	وحدة أولية للعينات	UPE
United Kingdom Data Archive	مركز أرشيف بيانات المملكة المتحدة تأويه جامعة إيسكس	UKDA
Universal Numeric Fingerprint	توقيع رقمي عالمي	UNF
	منظمة الأمم المتحدة للطفولة (اليونيسيف)	UNICEF
Uniform Resource Locator	سلسلة حروف تستعمل لتحديد البيانات على الأنترنت	URL
	ناقل متسلسل عام : Universal Serial Bus	USB
	لغة الترميز الموسعة	XML
	لغة هيكلية بيانات	XSL

مقدمة

مشاركة البيانات بطريقة أكثر منهجية ستعكس بفائدة على كل أنواع البحث عملياً [17].

أعد هذا الدليل لمساعدة منتجي البيانات الجزئية والمحتفظين بها على تشكيل سياساتهم وإجراءاتهم الخاصة بشأن بث ونشر ملفات البيانات الجزئية. من الضروري أن تحدد هذه السياسات والإجراءات بطريقة رسمية وشفافة، إذ لا يسمح بث البيانات على نحو ملائم بإتاحة البيانات وتوثيقها فحسب، بل يسمح أيضاً بصياغة الشروط التي تحكم الاستفادة من البيانات. إذ ينبغي أن تكون هذه المعلومات عامة ومتوافرة بسهولة، ويفضل أن يتم ذلك عبر شبكة الإنترنت.

وإذا كان القسم الأكبر من هذه الوثيقة يقدم سمة عامة، فهو موجه مع ذلك في المقام الأول إلى منتجي البيانات الرسمية - مثل مكاتب الإحصاءات الوطنية وما يشابهها من المؤسسات - في البلدان النامية. أما البيانات فعادة ما نعي بها البيانات الجزئية المستخرجة من الاستقصاءات عن طريق استطلاعات الرأي، والمسوح، وأنظمة جمع البيانات الإدارية.

أعد هذا الدليل برعاية الشبكة الدولية لاستقصاءات الأسر المعيشية IHSN. وهو يقوم بشكل أساسي على دراسات اللجنة الاقتصادية لأوروبا والأمم المتحدة، ولا سيما لجنة البحث التابعة لمؤتمر الإحصائيين الأوروبيين المكلفة بالتمهيد لإدارة السرية والنفاذ إلى البيانات الجزئية، ومكتب الإحصاءات في الاتحاد الأوروبي (Eurostat) [5] [24] [25] [26]. كما أنه يستند علاوة على ذلك إلى تجربة معاهد الإحصاءات القائمة في بلدان تطبق إتاحة ملفات البيانات الجزئية منذ زمن بعيد (منذ أكثر من 40 سنة)، وإلى تجربة مراكز بيانات جامعية عديدة.

تجيب المعلومات التي يحتوي عليها هذا المستند على إثني عشر سؤالاً يطرح في إطار وضع سياسة لنشر ملفات البيانات الجزئية.

1. ما هي البيانات الجزئية؟
2. ما هي البيانات الوصفية؟
3. ما هي الحجج التي تصب في صالح بث ونشر البيانات الجزئية؟
4. ما هي التكاليف والمخاطر المترتبة على بث البيانات الجزئية وكيف يمكن التحكم بها؟
5. إلى من تتوجه البيانات الجزئية؟
6. ما هي الشروط التي تحكم بث البيانات الجزئية؟
7. ماذا نعني بـ "إغفال" البيانات الجزئية؟
8. هل ينبغي أن يكون النفاذ إليها مدفوعاً أم مجانياً؟
9. في أي وقت من دورة بث البيانات الجزئية يجدر جعلها متاحة للعام؟
10. ما هي الشروط التي يجب استيفائها من حيث البنية التحتية التقنية؟
11. ما هي الشروط المؤسسية المتعلقة بث البيانات الجزئية؟
12. كيف يمكن الترويج لاستخدام البيانات الجزئية؟

يعالج هذا الدليل بشكل أساسي النواحي السياسية المتعلقة بث البيانات الجزئية، إذ يترتب على بث ملائم وآمن للبيانات الجزئية إيجاد حلول تقنية لغايات توثيق البيانات والبيانات الوصفية وإغفالها وفهرستها وحفظها. وقد تم التطرق إلى هذه المسائل بإيجاز هنا، غير أنها درست بالتفصيل في مستندات أخرى صادرة عن شبكة IHSN أو هيئات أخرى.

يعتبر جمع بيانات الإحصاءات المعدّة لدعم عمليات اتخاذ قرارات خاصة وعامة في بلد ما، مهمة كبيرة عادة ما تمولها الأموال العامة. لذا يتعين على جميع منتجي البيانات الممنوحة، والباحثين، والرعاة أن يحققوا نتيجة استثمار قصوى من خلال توظيف هذه البيانات.

إن البيانات الاجتماعية الاقتصادية التي تشكل المحور الرئيسي لهذه الوثيقة، هي ثمرة عمليات الاستقصاءات، والمسوح عن طريق استطلاعات الرأي وأنظمة التسجيل الإداري. ينتج عن هذه الأنشطة بيانات جزئية (وهي بيانات على مستوى الأفراد المستجيبين). ويمكن أن تتم لاحقاً معالجة هذه البيانات الجزئية (مراجعتها وتحليلها وتصنيفها) قبل أن توضع تحت تصرف المستخدمين. وعادة ما نحصل بهذه الطريقة على بيانات معتمدة تظهر في شكل جداول ورسوم بيانية وملفات وتقارير وصفية وتحليلات، يعتمد محتوى هذه الجداول والتقارير على أهميتها بالنسبة إلى منتجي البيانات والرعاة. تجرى معظم أنشطة جمع البيانات لغايات محددة، وتكون الأولوية المطلقة فيها عادة متمثلة في إرضاء المنتج أو الراعي. ومع ذلك، يمكن للبيانات المجموعة لغاية معينة أن تكون مفيدة لأشخاص آخرين كثر، لم يحسب لهم أي حساب لدى جمع البيانات. وبعبارة أخرى، يمكن لهذه البيانات أن تستخدم لغايات أخرى غير تلك التي جمعت لأجلها في الأساس. تشكل إتاحة هذه البيانات لجماعة الباحثين طريقة اقتصادية وفعالة لمضاعفة وتنويع طرق تحليل المعلومات المجموعة والاستفادة منها، فالاستفادة المعتمّقة من هذه البيانات توفر إمكانيات غير محدودة تقريباً للنفاذ إلى معلومات جديدة.

جعلت القوة المتنامية لأجهزة الكمبيوتر والبرمجيات منذ الثمانينات من القرن الماضي البيانات الجزئية أكثر جاذبية للباحثين. وأصبح منتج البيانات في جميع البلدان يواجهون طلباً متزايداً على النفاذ إلى البيانات الجزئية التي تمثل قاعدة للإحصاءات الرسمية. ولا يتيح النفاذ إلى البيانات الجزئية إجراء أبحاث جديدة وأكثر تنوعاً فحسب، بل يتيح أيضاً إمكانية تطوير طرق مبتكرة للاستفادة من المعلومات ومعالجتها ونشرها، ناهيك عن تشكيل مجموعات بيانات جديدة تمزج بين مصادر عدة .

ومع ذلك، يشكل اتخاذ القرار بشأن الطريقة الأفضل لنشر البيانات الجزئية تحدياً حقيقياً لمنهجها، وي طرح هذا التحدي من الناحية التقنية والتنظيمية، إذ تقضي العملية اعتماد إجراءات ملائمة لتوثيق البيانات الجزئية وفهرستها ومشاركتها. وقد وضعت منظمات الأرشيف والبيانات معايير دولية ومبادئ للممارسات السليمة في هذا الشأن من أجل هذه المسألة، ولكن التحدي ينطوي أيضاً على شق قانوني وأخلاقي. وإذا كان منتج البيانات على دراية تامة بقوة البيانات الجزئية وأهميتها، عليهم أن يوازنوا بين الطلب عليها وضرورة تأمين سرية المعلومات التي يقدمها الأشخاص المستجيبون. إذ ترعى هذا المطلب التشريعات الوطنية المتعلقة بالإحصاءات وبسرية المعلومات، وغالباً ما يشكل التزاماً تجاه المستجيبين أثناء جمع المعلومات. لذا يتعين على معاهد الإحصاء وغيرهم من منتجي البيانات أن يتوصلوا إلى الحفاظ على ثقة المستجيبين، تحت طائل تراجع الرغبة في التعاون لدى إجراء المسوح، وبالتالي تراجع جودة الإحصاءات. وفي هذا الإطار، يستوجب نشر البيانات الجزئية وضع سياسات وإجراءات تحدد بطريقة رسمية شروط النفاذ إليها.

يختلف الإطار المعدّ لهذه الغاية من بلد إلى آخر. ولكن يجدر التنويه بأنه "بغض النظر عن الاختلافات التي يمكن أن تتواجد بين الممارسات والسياسات بشأن مشاركة البيانات، والقيود القانونية التي يخضع لها النفاذ إلى البيانات، فإن

1. ما هي البيانات الجزئية؟

1.1 مفهوم البيانات الجزئية

على المدى الطويل، غير أنه لا يمكن فك تشفيرها أو استخدامها في قاموس للبيانات (ملف أو مستند منفصل). ويظهر الرسم 1 مقتطفاً من ملف نموذجي للبيانات الإحصائية بصيغة ASCII ثابتة.¹

لإنشاء جدول إحصائي أو تحليل بيانات ASCII، يجب في البداية استيراد هذه الأخيرة إلى برمجية أخرى. تقترح كل برمجيات الإحصاءات وقواعد البيانات أدوات وتعليمات لهذه الغاية. وتضم الصفحة القادمة مثالاً عن سكريبت `atS` يسمح باستيراد بيانات ASCII من الرسم 1 وبإقران ملصقات (labels) بالمتغيرات والرموز حتى يصبح تصفح المحتوى أكثر ملاءمة للمستخدم. ومن الواضح أن كتابة هذا النوع من السكريبت تتطلب من المستخدم أن يمتلك قاموساً للبيانات يصف المحتوى وبنية ملف البيانات من نوع ASCII.

عند استيراد ملف ASCII الظاهر في الرسم 1 إلى Stata بعد تنفيذ السكريبت، يظهر كما هو في مبيّن في الرسم 2. وتحتوي صيغ الملفات المملوكة SAS، أو SPSS، أو Stata أو البرمجيات المشابهة في آن معاً على البيانات والمتغيرات والملصقات الخاصة بها.

تسمح الاستقصاءات أو عمليات جمع البيانات الإدارية لمؤسسات الإحصاء أو غيرها من المؤسسات المنتجة للبيانات بالحصول على معلومات تخص الوحدات المدروسة: الأسر والأفراد والمؤسسات والشركات الزراعية والمدارس وهيئات الصحة وغيرها. ونعني بالبيانات الجزئية هنا ملفات البيانات الإلكترونية التي تحتوي على معلومات تخص كلاً من هذه الوحدات المدروسة. وفي هذا الصدد، تتعارض البيانات الجزئية مع البيانات الكلية التي تشكل ملخصاً عن المعلومات المجموعة بشكل معدلات أو نسب أو ترددات أو غيرها من إحصاءات الخلاصات.

كما تتشكّل البيانات الجزئية من ملفات بيانات، يضم كل خط (أو تسجيل) منها معلومات تخص الوحدة المدروسة. وتسجّل هذه المعلومات بشكل متغيرات من أنواع مختلفة (رقمية أو أوجدية، منفصلة أو مترابطة، إلخ). ويمكن الحصول عليها مباشرة من الشخص المستجيب، عن طريق نموذج استفتاء، أو عن طريق الملاحظة أو القياس (بواسطة نظام تحديد المواقع على سبيل المثال) أو الإلصاق أو الحساب.

إن المعلومات التي تحتوي عليها ملفات البيانات الجزئية الإحصائية مرمّزة. وبذلك، يمكن أن يكون جنس الشخص المستجيب مسجلاً بشكل متغيرة من نوع `Ho1a`. قد تعود إلى القيمة 1 أو 2 (1 للذكر، و2 للأنثى). وبالتالي، يجب مصاحبة البيانات الجزئية بقاموس بيانات يحتوي على قائمة بالمتغيرات المستخدمة، ووصف لمحتواها ومعنى كل رمز. وتشكّل هذه البيانات الوصفية الحد الأدنى من التوثيق المطلوب. غير أن عدد البيانات الوصفية الضرورية هو في الواقع أكبر بكثير، كما سنرى في الفصل الثاني.

عادة ما يضم مجموع بيانات المسوح أو الاستقصاءات ملفات عدة، غالباً ما تكون صادرة عن عدد من المستويات في عملية واحدة لجمع البيانات. وفي أغلب الأحيان، تسمح الاستقصاءات والمسوح التي تجري مع الأسر بجمع بيانات على مستويين على الأقل: الأسرة (مع متغيرات تصف ميزات المسكن مثلاً)، والفرد (مع معلومات عن السن والوضع المدني ومستوى التعليم والنشاط مثلاً). ويمكن تشكيل مجموعة من البيانات عن طريق ملف أو ملفات عدة لكل من هذين المستويين. ويضم هذان المستويان متغيرات أساسية تسمح للمستخدمين بتكوين علاقة بين المعلومات المتواجدة في الملفات المختلفة. ويدعى مجموع البيانات المنظمة بهذه الطريقة متسلسلاً هرمياً.

1.2 شكل تخزين ملفات البيانات الجزئية وبثها

يمكن تخزين ملفات البيانات الجزئية بصيغ عديدة. وتعدّ صيغة ASCII غير المملوكة والصيغ المملوكة التي ابتكرتها برمجيات الإحصاء المتخصصة، مثل SAS، وSPSS، وStata من بين الأكثر استخداماً. كما يمكن أن تحفظ البيانات الجزئية بصيغة SQL أو غيرها من صيغ البيانات الجزئية. غير أن هذه الصيغ أقل شيوعاً وأقل ملاءمة لبيانات المسوح والاستقصاءات، بما أن البرمجيات الأساسية للبيانات ليست مصممة خصيصاً لابتكار جداول وتحليلات إحصائية.

ليست صيغة الملف ASCII خاصة ببرمجية أو منصة معينة. إذ تحتوي ملفات ASCII على بيانات تستطيع معظم البرمجيات قراءتها. وبما أنها ليست مرتبطة ببرمجية قد تزول يوماً ما، تشكل الحل الأفضل لضمان حفظ البيانات

1 يمكن لملفات ASCII أن تكون ثابتة أو محددة، وتشغل البيانات المصاحبة لمتغيرة دائماً المكان نفسه (العمود نفسه). وفي ملفات ASCII المحددة، نجد المعلومات الخاصة بكل متغيرة منفصلة عن بعضها البعض بواسطة رمز مميز (فاصلة منقوطة، أو جدول، أو فاصلة أو غيرها من الرموز التي يحددها المستخدم). فعلى سبيل المثال، في ملف ASCII محدد بالفواصل، نجد كل متغيرة منفصلة عن الأخرى بواسطة فاصلة.

الرسم 1 مقتطف من ملف بيانات بصيغة ASCII ثابتة

العمود 1-3: متغيرة معرّف الأسرة
العمود 4: متغيرة المكان (الرمز = 2 ريفي)
العمود 5-6: متغيرة معرّف الشخص
العمود 7-8: متغيرة رابط القرابة برب الأسرة
العمود 9: متغيرة الجنس (1 = ذكر، 2 = أنثى)
العمود 10-11: متغيرة السن (بعدد السنوات)

التسجيل 1 (معلومات عن الشخص الأول)	→	12	1	114021						
التسجيل 2 (معلومات عن الشخص الثاني)	→	12	2	223921						
إلخ		12	3	321711						
		12	4	321311						
		12	5	32	5					
		12	6	31	1					
		22	1	124711						
		22	2	321611						
		22	3	814311						
		22	4	629933						
		32	1	113521						
		32	2	223922						
		32	3	31	1					
		32	4	1021612						
		32	5	102	4					
		32	6	101	4					
		41	1	117821						
		41	2	227521						

مثال عن صفحة عرض استيراد بيانات ASCII إلى Stata (set-up)

* . قراءة بيانات ASCII الموجودة في ملف test.dat واستيراد القيم بمتغيرات جديدة؛
.. test.dat باستخدام 10-11 age 9 sex 7-8 relat 5-6 pid 4 area 1-3 infix hhid ..

* . إضافة ملصق لوصف كل متغيرة جديدة؛

.. ملصق المتغيرة hhid "معرف الأسرة"؛

.. ملصق المتغيرة area "مكان السكن"؛

.. ملصق المتغيرة pid "معرف الشخص"؛

.. ملصق المتغيرة relat "رابط القرابة برب الأسرة"؛

.. ملصق المتغيرة sex "الجنس"؛

.. ملصق المتغيرة age "السن عند آخر عيد ميلاد"؛

* . إضافة ملصق لكل رمز تستخدمه المتغيرات؛

.. ملصق وصفي relatcod 1 "رب أسرة" 2 "زوج/ زوجة" 3 "ابن/ ابنة" 4 "ابنة الشريك/ ابن الشريك"

.. 5 "حفيد/ حفيدة" 6 "أب، أم" 7 "حمى/ حماة" 8 "أخ/ أخت" 9 "أهل آخرون"

.. 10 "لا صلة"، إضافة؛

.. ملصق يقيم relatcod relat؛

.. ملصق يصف areacod area 1 "حضري" 2 "قروي"، إضافة؛

.. ملصق يقيم areacod area؛

.. ملصق يصف sexcod 1 "ذكر" 2 "أنثى"، إضافة؛

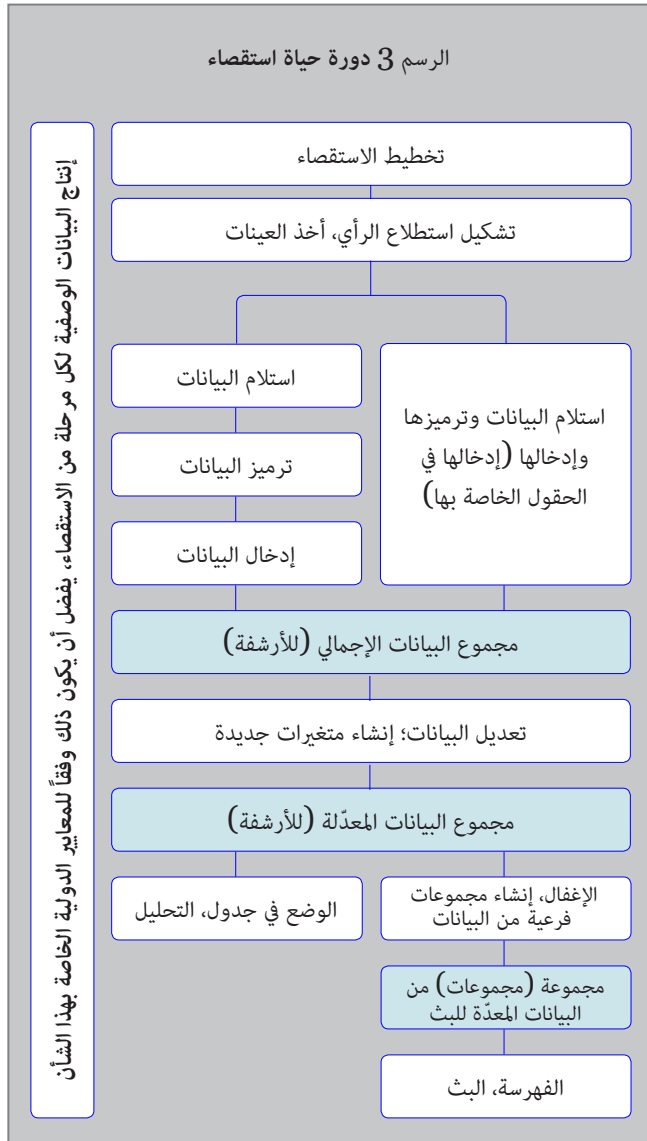
.. ملصق يقيم sexcod sex؛

* . حفظ الملف بصيغة ملف Stata؛

.. حفظ "test.dta"، استبدال؛

الرسم 2 مقتطف من ملف بيانات بصيغة Stata

hhid	area	pid	relat	sex	age	
40	ذكر	رب الأسرة	1	قروي	1	1
39	أنثى	زوج/ زوجة	2	قروي	1	2
17	أنثى	إبن/ إبنة	3	قروي	1	3
13	أنثى	إبن/ إبنة	4	قروي	1	4
5	أنثى	إبن/ إبنة	5	قروي	1	5
1	ذكر	إبن/ إبنة	6	قروي	1	6
47	أنثى	رب الأسرة	1	قروي	2	7
16	أنثى	إبن/ إبنة	2	قروي	2	8
43	ذكر	أخ/ أخت	3	قروي	2	9
99	أنثى	أب/ أم	4	قروي	2	10
35	ذكر	رب الأسرة	1	قروي	3	11
39	أنثى	زوج/ زوجة	2	قروي	3	12
1	ذكر	إبن/ إبنة	3	قروي	3	13
16	أنثى	لا صلة	4	قروي	3	14
4	أنثى	لا صلة	5	قروي	3	15
4	ذكر	لا صلة	6	قروي	3	16
78	ذكر	رب الأسرة	1	حضري	4	17
75	أنثى	زوج/ زوجة	2	حضري	4	18



ويقضي تعديل محتوى التسجيلات في ملفات البيانات الجزئية المعدلة للبث بمحو المعرّفات المباشرة وغير المباشرة، لحماية مجهولية هوية المستجيبين. ولا يترجم ذلك بالضرورة بمحو المتغيرات، بل يكفي أحياناً جمع متغيرات معينة في فئة تظهر تفاصيل أقل. ويمكن أن يمّر ذلك عن طريق اختزال عدد التسجيلات التي يحتويها ملف البيانات الجزئية المنشور (وهو حال بيانات تعداد السكان مثلاً). أما مفهوم الإغفال أو ضبط سرية الإحصائات فيقصد به عادة العمليات الرامية إلى حماية هوية المستجيبين.

تعدّ ملفات البيانات الجزئية المجموعة للإحصاءات الرسمية ملفات قابلة للبث، شرط ضمان سرية هوية المستجيبين بشكل صحيح وحمايتهم. وفي إطار تحديد سياسة النشر، ثمة ثلاثة أنواع محتملة من الملفات، وهي: الملفات المتاحة للاستخدام العام والملفات الخاضعة لترخيص والملفات التي يمكن النفاذ إليها في المراكز الأمنية حصراً. وتختلف هذه الملفات من ناحية سهولة النفاذ إليها ودرجة إغفال هوية المستجيبين.

يمكن لمجموع بيانات الاستقصاء أن يشمل مئات المتغيرات، وحتى الآلاف منها. لذا فإن كتابة السكريبت لاستيراد ملفات البيانات هذه وتوثيقها عن طريق ملف بصيغة ASCII تستغرق وقتاً وقد تكون كذلك مصدراً للأخطاء. ومن أجل تقليص المخاطر والاستخدام المريح، يتعيّن على منتجي البيانات أن يقدموا ملفاتهم، إما بصيغة ASCII بنماذج سكريبت SPSS، SAS وStata، وإما بالصيغ الإحصائية المملوكة الأكثر شيوعاً، وثمة برمجيات متخصصة مثل sfer-StatTran (Stata Corporation) لتحويل ملفات البيانات تلقائياً من صيغة رزمة إلى أخرى.

1.3 أي نسخة ينبغي بثها من ملفات البيانات؟

غالباً ما ينشئ منتجو البيانات نسخاً عديدة من ملف البيانات الجزئية نفسه. وتتميز هذه النسخ عن بعضها البعض من ناحية النوعية، والمحتوى، وعدد التسجيلات، وهي تتنوع بين ملفات البيانات الجزئية البحتة – التي تحتوي على كل الأجابة التي قدّمها كل مستجيب، والتي تم الحصول عليها مباشرة بعد إدخال البيانات – وبين ملفات للبث على نطاق واسع، منظفة ومعدّلة. يمثل الرسم 3 دورة حياة نموذجية لاستقصاء أو لمسح.

1.4 ما هي العناصر الحساسة في محتوى البيانات الجزئية؟

تستخدم البيانات المستخرجة من الاستقصاءات والمسوح عن طريق الاستفتاء حصراً لغايات الإحصاء أو البحث. ولأسباب عملية، يذكر عادة اسم المستجيب وعنوانه في استطلاع الرأي، لكنه نادراً ما يرد في ملفات البيانات. وكقاعدة عامة، لا تحتوي هذه الأخيرة إداً على مكونات حقيقية للمعرّفات المباشرة. وفي المقابل، تضم ملفات البيانات الإدارية بشكل متكرر الأسماء والعناوين وأرقام الهواتف وأرقام الضمان الاجتماعي، إلخ.

ومع ذلك، تضم معظم مجموعات البيانات معرّفات غير مباشرة. وهكذا، يمكن للموقع الجغرافي، وتأليف الأسرة (العمر والجنس)، والنشاط الاقتصادي أن يساعدوا في تحديد هوية المستجيبين.

تعدّ هذه المعلومات حساسة، لأنها قد تسمح بتحديد هوية الوحدات المدروسة. وثمة متغيرات أخرى حساسة نظراً إلى طبيعة المعلومات التي تحتوي عليها، مثل حالة الشخص الصحية وعاداته الجنسية ودخله، إلخ. وعلى سبيل المثال، تعدّ الاستقصاءات التي تجري لدى الشركات حساسة بطبيعتها، إذ يمكن أن تستخدم لغايات المنافسة.

1.5 ما هي أنواع ملفات البيانات الجزئية الرئيسية التي تبثّ؟

تختلف ملفات البيانات الجزئية المعدلة للبث دائماً تقريباً عن تلك المخصصة لاستخدام فريق عمل مؤسسات الإنتاج. وتتضمن عملية تحضير ملفات البيانات الجزئية الإجمالية التي سبثت عمليات تعديل للمحتوى و/ أو عدد التسجيلات.

على بيانات قد تدلّ على هوية الأشخاص إذا ما جمعت مع ملفات أخرى³.

يجب مسح المعرفات المباشرة مثل أسماء المستجيبين من مجموع البيانات الخاضعة لترخيص. غير أن هذه الملفات تحتوي أحياناً على متغيرات غير مباشرة قد تتيح التعرف إلى هوية المستجيبين حين تضمّ إلى مجموعات أخرى من البيانات (قوائم المصوّتين، السجلات العقارية، الملفات المدرسية على سبيل المثال).

يفضل أن يستند نشر الملفات الخاضعة لترخيص إلى صياغة وتوقيع عقد بين منتج البيانات والمستخدمين محل الثقة، أي أولئك الذين تعدّ حاجتهم إلى النفاذ إلى البيانات مشروعة، ويجب أن يذكر العقد شروط النفاذ إلى ملفات البيانات الجزئية واستخدامها، وتوقّع عقود الترخيص أحياناً من قبل المستخدمين الأعضاء في مؤسسة راعية ملائمة (مركز أبحاث، جامعة أو شريك في التنمية).

كما يوصى منتج البيانات بأن يستلموا من مقدمي الطلبات مسبقاً التوقيع على عقد للاستخدام والنفاذ، وأن يملأوا استمارة طلب تشرّح حاجة الرجوع إلى البيانات الخاضعة لترخيص (عوضاً عن النسخة المتاحة للاستخدام العام إذا وجدت)، وذلك لأغراض الأعمال الإحصائية أو لأغراض البحث. ويقدم الفصل السادس نموذجاً عن عقد واستمارة طلب خاصين بالملفات الخاضعة لترخيص. وتسمح هاتان الوثيقتان بفهم شروط منح ترخيص النفاذ إلى ملفات البيانات الجزئية.

البيانات المتاحة للقراءة في مراكز البيانات المحمية

يمكن اقتراح قراءة بيانات معينة في مراكز محمية (أو "حاويات بيانات")، في ظل شروط صارمة للغاية. أما هذه المراكز فهي كتابة عن قسم مجهز بأجهزة حاسوب غير متصلة بشبكة الإنترنت، ولا بشبكة خارجية، ولا يمكن تحميل أي معلومة منها على مفتاح USB، أو أقراص مدمجة أو أقراص DVD، أو أي نوع آخر من القارئات. تضم هذه "الحاويات" بيانات حساسة للغاية أو بيانات تتيح مباشرة وبسهولة تحديد هوية المستجيبين.

قد يتعلق الأمر مثلاً بمجموعات بيانات كاملة ناتجة عن المسوح والاستقصاءات التي أجريت في الشركات أو حتى الملفات الصحية التي تحتوي على معلومات في غاية السرية. أما المستخدمون الذين يؤدّون قراءة مجموعات البيانات هذه، فلا يستطيعون بالضرورة النفاذ إليها بالكامل، بل بطريقة محدودة فقط (النفاذ اللازم إلى المجموعات الفرعية من البيانات). وهم مدعوون لملء استمارة طلب تبيّن أن النفاذ إلى البيانات يخدم غايات إحصائية أو غايات أبحاث (أنظر المثال في الفصل السادس). كما يجب أن تخضع النتائج التي يحصلون عليها لفحص دقيق في إطار عملية كاملة لضبط السرية.

إن لاستخدام مركز بيانات محمي تكلفة لا يستهان بها. إذ يتطلب ترتيب المباني وشراء تجهيزات معلوماتية مخصصة. كما يتطلب وجود فريق عمل يتمتع بالكفاءات ويمكّن الوقت اللازم للتحقق من حسن سير العملية لتقليل خطر إفشاء السرية. ويجب أن يكون الأشخاص المعنيون مطلعين على طرق

"لا يحق لأي شخص (... أن يدعي تمتعه بحق الحصول على، أو النفاذ إلى بيانات يمكن التعرف من خلالها إلى هوية الأشخاص المشاركين فيها (...). بفضل مركزه المهني. فالنفاذ إلى هذا النوع من البيانات لا يحدّد فقط وفقاً للمركز المهني، بفضل التبعية لمؤسسة أو الالتزام المالي. فالأهم من ذلك هو الحاجة إلى هذه البيانات، والغاية من استخدامها، ودور المطالب ومسؤوليته في مسألة جميع البيانات. وبما أن النفاذ إلى البيانات التي تتيح التعرف إلى هوية الأشخاص المشاركين فيها لا يخلو من المخاطر، فهو يخضع إلى تقييم ومتابعة معتمّين بعد منح الرخصة المتعلقة بهذا الشأن". [14]

ملفات البيانات الجزئية المتاحة للاستخدام العام

يمكن لأي شخص يقبل بالخضوع لبعض الشروط الأساسية السهلة تصفّح ملفات البيانات الجزئية المنشورة على نطاق واسع. وهي شروط للاستعمال (مثل منع المتاجرة بالبيانات) وليس شروطاً على النفاذ إلى البيانات. تنشر بعض ملفات البيانات المتاحة للاستخدام العام من دون شروط؛ وهي غالباً ما تكون متاحة على شبكة الإنترنت. يمكن النفاذ إلى هذه البيانات بسهولة، لأن خطر التعرف على هوية المستجيبين يعتبر ضئيلاً للغاية. يعني ذلك حذف كامل المحتوى الذي قد يسمح بتحديد هوية المستجيبين مباشرة، كالأسماء والعناوين وأرقام الهاتف بشكل خاص. كما ينبغي التخلّص من المعرفات غير المباشرة أيضاً. وتختلف هذه المعرفات بحسب منهجية الاستقصاء، لكن بعض عناصر المعلومات تحذف عادة، مثل المعلومات الجغرافية الأكثر تحديداً من منطقة العينات. كما تحذف تسجيلات معينة من البيانات المتاحة للاستخدام العام أحياناً، مثل المتغيرات المتطرفة أو المتسمة بتوزيع غير متماثل للغاية. ومع ذلك، ثمة طرق أخرى لضبط سرية الإحصاءات تسمح بتقليل خطر الكشف عن هوية الأشخاص، وتوفّر أكبر قدر من محتوى المعلومات (إعادة تنظيم القيم المتطرفة العليا والدنيا، أو حذف المتغيرات الواردة من مستجيبين معينين، أو حتى تقنيات تشويش البيانات²). غالباً ما تنشأ ملفات البيانات المتاحة للاستخدام العام عن طريق ملفات بيانات مسوح (وفي معظم الأحيان تكون مجموعات فرعية من التسجيلات وليست ملفات كاملة) واستقصاءات أجريت مع الأسر. ومع أن إنشاء ملفات بيانات متاحة للاستخدام العام عن طريق استقصاءات تجرى مع الشركات أمر ممكن من الناحية التقنية، إلا أنه يشكل تحديات سنتحدث عنها بالتفصيل.

يجب أن تقدّم ملفات البيانات المتاحة للاستخدام العام القدر الأكبر من المعلومات. ووفقاً لمركز إحصاءات الصحة الأمريكي NCHS، "فإن الغاية تتمثل في إتاحة بيانات جزئية على أوسع نطاق وبأكبر تفصيل ممكن، وذلك في حدود المصادر المتاحة، والمستلزمات النوعية، والقيود التقنية وضرورة حماية سرية البيانات فقط" [14] (ترجمة من الإنجليزية)

الملفات الخاضعة لترخيص

تتميز الملفات الخاضعة لترخيص، أو ملفات الأبحاث، عن الملفات المتاحة للاستخدام العام بأنها تنشر حصراً على المستخدمين الذين يحملون ترخيصاً يخولهم النفاذ إليها، وهو يمنح لهم بعد تقديم طلب مبرر وفق الأصول وتوقيع اتفاق يحكم استخدام هذه البيانات. وكقاعدة عامة، تغفل هوية المستجيبين في الملفات الخاضعة لترخيص من أجل تقليل خطر التعرف إلى الأشخاص إلى أقصى حد في حال استخدام منفصل. غير أنها يمكن أن تحتوي

3 للحصول على معلومات أكثر تفصيلاً بشأن تعريف ملفات البيانات المتاحة للاستخدام العام والملفات الخاضعة لترخيص واختلافاتها، يرجى قراءة العمل الذي أعدته اللجنة الاقتصادية لأوروبا والأمم المتحدة في إطار مؤتمر الإحصائيين الأوروبيين (24).

2 يقدم الفصل السابع استعراضاً ووصفاً لتقنيات ضبط سرية الإحصاءات الرامية إلى تقليل خطر الإفصاح عن هوية المستجيبين.

وإدارتها، وإليك بعض الأمثلة مع ذكر عنوان الموقع الإلكتروني الخاص بها: مركز MCRDC لأبحاث المسوح، وهو مشروع مشترك بين مكتب reau-Census Bu الأمريكي وجامعة ميتشيجان (www.isr.umich.edu/src/merdc/); ومركز ناشيونال أوبنيون للأبحاث (NORC)

تحليل البيانات، وملمين بطرق معالجة الطلبات وتنظيم خوادم الملفات.

ونظراً إلى تكلفة الاستخدام الكبيرة والمهارات التقنية العالية المطلوبة، اختارت بعض معاهد الإحصاءات وغيرها من منتجي البيانات الرسمية أن تتعاون مع هيئات جامعية أو مراكز أبحاث لإنشاء "حاويات بيانات"

الإطار 1 استقصاء واحد، ومنتجات عديدة

دقة، ولكن مميزات أقل تفصيلاً (ملفات الدولة؛ نسبة 5% لأخذ العينات). (ترجمة من الإنجليزية)

المصدر: <http://www.census.gov/population/www/cen2000/pums/index.html>

تمت زيارة الموقع في 8 أبريل/ نيسان 2010

ويمكن مراجعة كامل البيانات الجزئية في مختلف مراكز البيانات المحمية في الاتحاد الأوروبي، لا سيما في مركز MCRDC التابع لجامعة ميتشيجان.

يسمح مركز أبحاث المسوح في ميتشيجان (MCRDC) للباحثين المؤهلين العاملين على مشاريع معتمدة من قبل مكتب Census Bureau الأمريكي باستخدام البيانات غير المنشورة المجموعة في إطار البرامج الاقتصادية والديموغرافية التي يجريها Bureau Census، والمركز الوطني للإحصاءات الصحية (NCHS). وتجري كل أبحاث مركز MCRDC في قلب مختبره المحمي، الواقع في آن هاربور في قلب معهد الأبحاث الاجتماعية التابع لجامعة ميتشيجان. (ترجمة من الإنجليزية)

المصدر: <http://www.icpsr.umich.edu/icpsrweb/ICPSR/access/restricted/index.jsp>

(تمت زيارة الموقع في 7 مايو/ أيار 2010)

يمكن لمنتجي البيانات أن يختاروا إصدار منتجات عديدة عن طريق مجموعة واحدة من بيانات المسوح أو الاستقصاءات. وهكذا يمكن أن تنشأ ملفات البيانات الجزئية المتاحة للاستخدام العام عن طريق عينات صغيرة أو مجموعات فرعية من المتغيرات. ويمكن أن تنشر هذه الملفات على نطاق واسع من دون خطر إفشاء هوية المستجيبين. كما يمكن توفير نسخة موسعة أكثر تتناول عينة أكبر ويكون النفاذ إليها مشروطاً بالحصول على ترخيص. وأخيراً، يمكن قراءة الملف الكامل (مع المعرفات أو من دونها) في مركز محمي.

أنج مكتب Census Bureau الأمريكي بهذه الطريقة ملفين مختلفين متاحين للاستخدام العام انطلاقاً من مجموع البيانات الناتجة عن مسوحات عام 2000، وأخذت عينات لهذه الغاية بنسبة 1% و 5%.

نظراً إلى التطور السريع الذي تشهده تقنية المعلوماتية وقدرة النفاذ المتنامية إلى بيانات المسوح التي تتمتع بها جماعة المستخدمين، اضطر مكتب Census Bureau إلى اعتماد تدابير أكثر صرامة لحماية سرية البيانات الجزئية المتاحة للاستخدام العام، وذلك باتباع تقنيات تحدّ من إفشاء المعلومات. ويقرّ مكتب Census Bureau من جهة أخرى، أن المستخدمين يحتاجون إلى بيانات أكثر تفصيلاً تكون محصورة أكثر من الناحية الجغرافية. لذا ستتوفر مجموعتان من الملفات: تضمّ الأولى عدداً أكبر من المميزات المفصلة (ملف وطني، نسبة 1% لأخذ العينات)؛ في حين تضم الثانية بيانات جغرافية أكثر

إنشاء عملية تسمح لهم بأن يقدموا برامج معالجة البيانات وتحليلها عن بعد. ويعطى المستخدم مجموعة من بيانات مرّبة تعيد إنتاج بنية البيانات الأصلية ومحتواها، وفي هذا الإطار، يستطيع الباحثون بلورة برامج مساعدة أدوات مثل SAS، أو SPSS أو Stata. ثم تنقل هذه البرامج إلى فريق عمل مودع البيانات، الذي ينفذ التطبيق على مجموع البيانات الأصلي. أما نتائج هذه العملية، فيتم التحقق منها وتعديلها قبل إرجاعها إلى المستخدم.

نذكر على سبيل المثال مشروع دراسة الدخل في لوكسمبورغ LIS الذي يتيح إمكانية النفاذ إلى قواعد البيانات الجزئية عن طريق نظام تقديم أعمال يدعى LISSY (الإطار 2).

لا شك أن سرية البيانات محمية، ولكن يمكن لتكلفة تقديم الأعمال أن تكون مرتفعة. بالإضافة إلى ذلك، إذا لم تمنح الموارد الكافية، ستكون العملية بطيئة بالنسبة للمستخدمين.

النفاذ إلى البيانات عن بعد

يتمتع المستخدمون هنا بحق النفاذ إلى برمجيات تصنيف البيانات وتحليلها،

التابع لجامعة شيكاغو (www.norc.org/DataEnclave)؛ برنامج مركز بيانات البحوث التابع لإحصاءات كندا Statistique Canada (www.statcan.gc.ca/rdc-cdr/index-fra.htm)؛ ومركز أبحاث البيانات التابع لمركز إحصاءات الصحة الأمريكي (www.cdc.gov/nchs). (http://www.cdc.gov/nchs)

1.6 هل من بدائل لمشاركة البيانات الجزئية؟

في الحالات المذكورة أعلاه، يقدّم للمستخدم حق النفاذ المباشر إلى ملفات البيانات الجزئية. ولكن ثمة نوع آخر من أشكال النفاذ إلى البيانات الجزئية، لعلّ أبرزها تقديم الأعمال عن بعد والنفاذ عن بعد إلى البيانات الجزئية. أما الشروط الخاصة بالسياسات التي تحكم هذا النوع من النفاذ فهي غير مفصلة هنا، بل مذكورة فحسب، مع إشارات إلى مصادر المعلومات المكتملة. ويجدر بالذكر أن هذه الأنظمة عادة ما تكون مكلفة ومعقدة تقنياً.

تقديم الأعمال

تتمثل إحدى الوسائل المتاحة للمستخدمين لغاية تحليل البيانات السرية في

يناسب هذا النظام إنشاء الجداول الإحصائية (وهي حالة مسح السكان والمسكن بشكل خاص)، بل يخدم أيضاً التحليل المعمق.

توظيف باحث جوجل مؤقت

يحرص بعض منتجي البيانات على أن يتمكن الباحثون من النفاذ إلى البيانات الجزئية عن طريق توظيفهم محرك البحث جوجل. وهم يخضعون إزاءً إلى نفس القواعد المتعلقة بالسرية التي يخضع لها فريق العمل الدائم في المؤسسة. ولا يتم اللجوء إلى هذه الوسيلة إلا إذا كان نشاط الباحث يشكل مساهمة حقيقية في عمل منتج البيانات المعني، تحت طائل المحاكمة الوهمية البسيطة. [24]

غير أنهم لا يستطيعون تحميل أي مجموعة من البيانات، ولا إنشاء جداول تكشف عن المعلومات الفردية أو عدد قليل من التسجيلات.

نجد في السوق برمجيات عدة (ross, Beyond 20/20, Nesstar, PcAxis, Redatam, SuperC، إلخ). وتطوّر بعض المراكز الفائزة التطور مناصتها الخاصة، إذ يدير مركز الأرشفة UKDA قسماً للبيانات الآمنة (SDS) ”يهدف إلى تعزيز امتياز الأبحاث عن طريق توفير إمكانية النفاذ الآمن عن بعد للباحثين الأهل للثقة إلى البيانات التي تعدّ في غاية الحساسية أو التفصيل أو السرية أو تلك التي يسهل التعرّف من خلالها إلى المستجيبين، وذلك بوضعها في تصرفهم بواسطة عقد ترخيص أو تدابير نشر معيارية“، [12]

الإطار 2 الإطار 2دراسة الدخل في لوكسمبورغ LISSY - Luxembourg Income Study

ويؤدي النظامان إلى النتائج نفسها، غير أن موقع LIS ينصح المستخدمين بشدة بالمرور عن طريق JSI للنفاذ إلى LISSY. فاستخدام هذه الواجهة أسهل، وهي تعرض خيارات إضافية مثل النفاذ إلى تاريخ مهمات المستخدم.

تعليمات التقديم

أياً كانت طريقة تقديم الأعمال، فثمة بعض المزايا المتعلقة بتركيب البرمجة المعيارية التي ينبغي أن تتقن حتى يتمكن نظام LISSY من معالجة طلبات المستخدمين على نحو صحيح.

أما برمجيات الإحصاء المتاحة حالياً في نظام LISSY فهي SAS 9.2، و Stata 11.0 و SPSS 11.5.

المساعدة التقنية

ينبغي توجيه كل الأسئلة المتعلقة باستخدام قواعد بيانات LIS ومحتواها إلى قسم المساعدة التقنية (LIS User Support) وليس إلى أعضاء فريق عمل LIS. أما الهدف فيتمثل في جرد مجموع الأسئلة المطروحة بطريقة متناسقة.

LISSY هو نظام تنفيذ مهمات عن بعد بالكامل يعمل على مدار الساعة طوال أيام الأسبوع. وهو يسمح للباحثين بتقديم برامج معالجة إحصائية على دفعات (بواسطة برامج SAS، أو SPSS أو Stata) من دون أن يتنقلوا. وينفذ نظام LISSY تلقائياً المهام المطلوبة، ثم ينقل النتائج بشكل مركّب، بمعدل يضع دقائق.

لا يمكن تحميل قواعد البيانات الجزئية، كما لا يمكن النفاذ مباشرة بأي شكل من الأشكال إلى البيانات. إنما ترسل إلى المستخدمين نتائج الطلب الإحصائي فقط.

التسجيل إلزامي

توضع الأرقام الأساسية لنظام LIS (Key figures) تحت تصرف العامة، لكن النفاذ إلى قواعد البيانات الجزئية محصور على المستخدمين المسجلين. ويمنح الترخيص بالنفاذ لفترة سنة واحدة فقط، قابلة للتجديد كل عام.

طريقتان لتقديم الأعمال

يقدم LISSY نفاذاً آمناً عن بعد إلى البيانات الجزئية عن طريق وسيلتين للتقديم، هما:

• برمجية ISJ (Job Submission Interface)

• برمجية البريد مثل Outlook، Thunderbird، إلخ

المصدر: <http://www.lisproject.org/data access/data access.html> (تمت زيارة الموقع بتاريخ 9 إبريل/ نيسان 2010).

2. ما هي البيانات الوصفية؟

تعرف البيانات الوصفية بشكل عام على أنها "بيانات تصف بيانات أخرى". ويذكر الفصل السابق الأهمية التي تتسم بها بها عملية إتاحة معجم ملائم للبيانات يصف محتوى كل المتغيرات التي تضمها مجموعة معينة من البيانات. لكن البيانات الوصفية تقدّم معلومات تفوق بكثير البيانات التي يقدّمها معجم بيانات بسيط.

تهدف البيانات الوصفية إلى مساعدة الباحثين في فهم ما تقيسه البيانات والطريقة التي تم الحصول عليها بها. فمن دون وصف جيد لمفهوم الاستقصاء والطرق المستخدمة في جمع البيانات ومعالجتها، يرجح أن يسيئ المستخدم تحليل هذه الأخيرة، وحتى أن يستخدمها بغير حكمة.

يقلص التوثيق الملائم حجم عمل الإحصائيين من حيث المساعدة التقنية الموجهة إلى المستخدمين الخارجيين لبياناتهم الجزئية.

ومن ناحية أخرى، تهدف البيانات الوصفية إلى مساعدة المستخدمين في تقييم جودة البيانات. فمن الضروري لكل باحث يرغب في الحكم على أهمية بيانات معينة لعمله أن يعرف المعايير النافذة فيما يتعلق بجمع البيانات، وأن يكون قادراً على تحديد الانحراف عن هذه المعايير.

وأخيراً، تعدّ البيانات الوصفية مهمة لتطوير أدوات البحث عن البيانات، مثل فهارس الاستقصاءات، التي تساعد في العثور على مجموعات من بيانات ذات صلة بموضوع البحث.

وينبغي التشديد على أن البيانات يجب أن تكون موثقة حتى وإن لم تكن معدّة للنشر. إذ يسهّل إنتاج بيانات وصفية ذات جودة بناء "ذاكرة مؤسسية" من حيث جمع البيانات، كما يمكن أن يساهم في تدريب فريق عمل جديد وتحسين تماسك البيانات مع الوقت.

ما هي البيانات الوصفية ذات الجودة؟

إن التعريف التالي للبيانات الوصفية ذات الجودة مقتطف من دليل entation- Good Practices in Data Docum (الممارسات السليمة في توثيق البيانات) الذي نشره مركز UKDA. [20] كما أن الموقع الإلكتروني للشبكة الدولية لمسوح الأسر (www.ihsn.org) ودليلها العملي الذي يتناول مسألة أرشيفات البيانات (reference Guide for Data Archivists) (Quick) يضمن أيضاً معلومات مهمة للغاية حول هذا الموضوع. (4)

"إن إنشاء مجموعة بيانات ذات جودة وقابلة للاستخدام على المدى الطويل يقتضي خطوة حاسمة تتمثل في جعل البيانات سهلة الفهم والتحليل. ويقتضي ذلك وصف البيانات وتوثيقها بطريقة سهلة وواضحة ومفصلة وشاملة في الوقت عينه".

(http://www.data archive.ac.uk)

ويشمل التوثيق المثالي لمجموعة بيانات بشكل أساسي ثلاثة أنواع من الوثائق:

1. الوثائق التفسيرية

تمثل الحد الأدنى المطلوب لضمان جدوى البيانات وإمكانية توظيفها على المدى الطويل - ومن دون هذه المعلومات، يستحيل فهم مجموع البيانات ومحتواها بشكل سليم.

معلومات حول طرق جمع البيانات

خصص هذا القسم لعملية جمع البيانات - الاستقصاء، جمع البيانات الإدارية أو نسخ وثيقة تمثل مصدراً. وهو يعرف بالأدوات المستخدمة، والوسائل المتبعة، والطريقة التي تبلورت فيها هذه الوسائل. وعند الاقتضاء، تقدّم تفاصيل حول طريقة أخذ العينات وإطار هذه العملية. بالإضافة إلى ذلك، ثمة معلومات في غاية الأهمية تتعلق بنظام جديد للإشراف على جمع البيانات وضبط الجودة. معلومات حول هيكلية مجموع البيانات

تعدّ في الأساس وثيقة مفصلة تصف هيكلية مجموع البيانات، بما فيها العلاقات بين بعض الملفات أو عناصر المعلومات (التسجيلات) التي تحتوي عليها. وهي تحدد بشكل عام المتغيرات الأساسية اللازمة لتحديد المواضيع في الملفات المختلفة بطريقة أحادية. وعادة ما يُشترط عدد الحالات والمتغيرات التي يحتوي عليها كل ملف، بالإضافة إلى العدد الإجمالي للملفات التي تشكل مجموع البيانات. وفي حالة النماذج الترابطية، يجب تقديم المعلومات المتعلقة ببنية التسجيلات والعلاقات بينها.

المعلومات التقنية

نخص هذه المعلومات البنية التحتية التقنية وهي تحدد عادة:

- النظام المعلوماتي المستخدم لإنشاء الملفات؛
- البرمجيات المستخدمة لإنشاء الملفات؛
- طريقة تخزين البيانات؛
- القائمة الكاملة بالملفات التي يحتوي عليها مجموع البيانات.

المتغيرات والقيم، أنظمة الترميز والتصنيف

يفضّل أن يشمل التوثيق قائمة كاملة بالمتغيرات (أو الحقول) التي تظهر في مجموع البيانات، مع وصف شامل ومعلومات مفصلة عن أنظمة الترميز والتصنيف المطبقة على المعلومات الخاصة بها. ومن المهم للغاية أن تذكر الحقول غير المملوءة والناقصة وأن تؤخذ في الاعتبار. ومن المفيد أيضاً تحديد المتغيرات التي لا تطبق عليها أي تسمية معيارية، وتحديد النسخة المستخدمة من نظام التصنيف، كما يفضّل أن تذكر المراجع الخاصة بها.

معلومات حول المتغيرات المشتقة

ينشئ منتجو بيانات عديدون متغيرات جديدة عن طريقة البيانات الأصلية. وقد تكون الطريقة بسيطة للغاية وتترجم مثلاً في إعادة تنظيم للبيانات وفقاً للسن (بعدد السنوات المنقضية)، بحسب فئات الأعمار ذات الصلة بالاستقصاء المعني. أما الطرق الأكثر تعقيداً فتتطلب معادلات حسابية متقدمة للغاية. وينبغي تفسير المنطق الذي يحكم إعادة التنظيم أو اشتقاق المتغيرات. يمكن إعادة تنظيم بسيطة وفقاً للسن أو توضّح في معجم البيانات. أما الاشتقاق الأكثر تعقيداً

- المنشورات التي ساهمت فيها الدراسة أو السياسات التي وضعت استجابة لها؛
- أي معلومة أخرى تعد ذات صلة بالموضوع.

مصدر مجموع البيانات

ينبغي تقديم معلومات تتعلق بعملية جمع البيانات، والتغيرات التي أجريت عليها، وتطور البيانات نفسها، والمنهجية والتعديلات المطبقة. ويمكن تقديم المعلومات التالية أيضاً:

- قوائم البيانات المغلوطة؛
- المشاكل التي تمت مواجهتها أثناء جمع البيانات وإدخالها وضبطها وتنظيفها؛
- التحويل الرامي إلى تحقيق التوافق مع برمجة أخرى أو نظام تشغيل آخر؛
- مراجع التقارير أو المنشورات التي استندت إلى الدراسة؛
- أي معلومة أخرى تعد ذات صلة بدورة حياة مجموع البيانات.

مجموع البيانات المتسلسلة زمنياً والثابتة، النسخ الجديدة

في حالة البيانات المتسلسلة زمنياً، أو البيانات الطولية، أو البيانات المستعرضة الثابتة، من غير المجدي نهائياً الاستفادة من المعلومات الإضافية، مثل تلك المتعلقة بتطور صياغة الأسئلة، أو عدم جدوى المتغيرات، أو إجراءات أخذ البيانات.

3. أنظمة الفهرسة

لهذه الأنظمة غاية مزدوجة. فهي تشكل أولاً مراجع لمجموع البيانات، ما يسمح بتحديدتها وذكرها بشكل صحيح في المنشورات. كما تسمح بتسجيلها رسمياً لغرض حفظها على المدى الطويل. وثانياً، تشكل الأداة الأساسية المستخدمة في البحث عن المصادر. وهكذا يمكن تحديد مجموع البيانات من غيره بطريقة أحادية، وتحديد المعلومات التي ستساعد المستخدمين الثانويين في التعرف على جدوى الدراسة بالنسبة إلى نشاطهم.

من دون العناوين والملخصات والكلمات المفاتيح وغيرها من العناصر المهمة في البيانات الوصفية، سيصعب على الباحثين تحديد مجموع البيانات والمتغيرات التي تخدم حاجاتهم. إذ تستند كل أنظمة الفهرسة والبحث عن المصادر، سواء أكانت مادية أم إلكترونية، على بيانات وصفية.

”يكون الحصول على توثيق ذي جودة أسهل حين يخطط لبنائه منذ بداية المشروع ويتمّ إعداده في كل مراحل البحث (على طول دورة حياة البيانات). ويمكن للتخطيط المفصل أن يقلص إلى حد بعيد الوقت والميزانية اللازمين للتخضير للتوثيق“ (http://www.data-archive.ac.uk). أنظر أيضاً [21].

المعايير والممارسات السليمة المتعلقة بالبيانات الوصفية

”إن التوافق التقني والدلالي ضروري لتسهيل وتشجيع النفاذ إلى بيانات البحث واستخدامها في سياق دولي تتشابه فيه الاختصاصات. وينبغي لتدابير النفاذ أن تأخذ بالاعتبار على نحو قانوني المعايير الدولية الثابتة المطبقة في مسألة توثيق البيانات“ (17)

فيجب أن تخضع لتوثيق منفصل، ويفضل أن يضم التوثيق هياكل تنظيمية أو تفسيرات منطقية دقيقة. وتكمن الفكرة في أن تتيح المعلومات المقدمة إمكانية معرفة العلاقة بين المتغيرات الأساسية وتلك الناتجة بسهولة. كما يوصى بحفظ الهياكل التنظيمية المعلوماتية المستخدمة في إنشاء المتغيرات المشتقة وتسجيلها إلى جانب المعلومات المتعلقة بالبرمجية المستخدمة.

الترجيح والاستكمال الخارجي

يجب تقديم قائمة كاملة بالمتغيرات المرجحة والمستكملة خارجياً، مع تعليمات تتعلق بطريقة الحصول على هذه المتغيرات وتذكر بوضوح شروط التطبيق التي تسري عليها. وتبرز أهمية ذلك حين يتعين تطبيق عوامل ترجيح مختلفة وفقاً للغاية المرجوة.

مصدر البيانات

عادة ما تقدّم معلومات مفصلة حول مصدر البيانات. وحين يكون مصدر البيانات مؤلفاً من إجابات على نماذج استطلاعات رأي على سبيل المثال، يجب أن تدخل كل إجابة بدقة في قوائم التوثيق. وفي الحالة المثالية، تذكر مراجع كل متغيرة. كما يجدر تفسير الظروف التي طرحت فيها الأسئلة، وإذ أمكن، ذكر الحالات التي تطبق فيها، ويفضل أن تقدّم بالإضافة إلى ذلك خلاصة بإحصاءات الإجابات.

السرية والإغفال

من الضروري تحديد ما إذا كانت البيانات تحتوي على معلومات سرية تتعلق بالأشخاص أو الأسر أو المؤسسات. وفي هذه الحالة، يوصى بذكر هذه البيانات، وكل عقد يحكم شروط استخدامها (مثل معلومات المشاركين في الاستقصاء). ويمكن لمسائل السرية أن تحدّ من تحليل البيانات أو نشر النتائج، لا سيما إذا كانت البيانات موجهة لاستخدام ثانوي. وفي حال إغفال البيانات من أجل حماية هوية الأشخاص المستجيبين، ينبغي تحديد الإجراء المتبع في هذا الإطار وأثره على البيانات. وإذا كانت هذه التغيرات قادرة على الحد من التحليل اللاحق، من المفيد ذكر ذلك.

2. المعلومات السياقية

ينبغي تقديم معلومات إلى المستخدمين تتعلق بسياق جمع البيانات، والاستخدام الذي وُجّهت له، وتغني هذه المعلومات التوثيق وتعمّقه. كما تسمح للمستخدمين الثانويين بفهم السياق الكامل لجمع البيانات وإجراءات جمعها. وهكذا تشكل شهادة تاريخية أساسية للباحثين في المستقبل.

وصف المشروع الأصلي

يفضل أن يشمل التوثيق معلومات عن تاريخ المشروع أو العملية التي نتجت عنها مجموعة البيانات في الأساس، والإطار الفكري والشكلي الأصلي. ويضم ذلك على سبيل المثال:

- الأسباب التي استدعت جمع البيانات؛
- أهداف المشروع وغاياته؛
- موضوع الدراسة؛
- المنطقة الجغرافية والإطار الزمني؛

عادة ما يصدر منتجو البيانات معاجم لرموز صيغة النص. ومن أجل توظيف تكنولوجيا الإنترنت على أفضل وجه، أصبحت معظم المعايير تحدد بلغة XML. أما مبادرة توثيق البيانات Documentation Initiative (DDI) Data فهي معيار خاص بتوثيق البيانات الجزئية⁴. [13]

سمحت مبادرة DDI بوضع معايير وفرت إطاراً تنظيمياً لترتيب محتوى البيانات الوصفية وتقديمها ونقلها وحفظها، وذلك في مجال العلوم الاجتماعية والسلوكية، ويسمح هذا الإطار بتوثيق ملفات البيانات الجزئية الأكثر تعقيداً بطريقة سلسلة ودقيقة في آن واحد.

تهدف مبادرة DDI إلى وضع معيار دولي يقوم على لغة XML لتوثيق البيانات الجزئية. أما الغاية فتتمثل في طرح وسائل عملية لتسجيل النواحي الأساسية من مجموع البيانات الجزئية وتبادلها. وتعدّ مبادرة DDI تطوراً هائلاً لمعجم الرموز الإلكترونية التقليدي؛ ففي الجوهر، تقدّم الإمكانات نفسها، غير أنها تحسّن إلى حد كبير نطاق المعلومات المقدّمة ودقّتها. وقد أطلق مبادرة DDI للبيانات الوصفية اتحاد أبحاث العلوم السياسية والاجتماعية بين الجامعات ICPSR، وهي منظمة تضم أكثر من 500 جامعة ومؤسسة تعليمية عليا من كل أنحاء العالم. وقد تولى المشروع اتحاد مؤسسات شمال أمريكية وأوروبية، يعدّ أعضاؤها من بين أهم منتجي البيانات ومؤسسات الأرشيف في العالم.

أنشئت مبادرة DDI لتعالج جميع أنواع البيانات الناتجة عن استقصاءات

طور اتحاد الأرشيفات مجموعة معايير خاصة بالبيانات الوصفية، وذلك حرصاً منه على تشجيع تبادل البيانات بين المنظمات وأنظمة البرمجيات وتحسين جودة التوثيق الإحصائي المقدّم للمستخدمين. وتقدّم هذه المعايير إطاراً منظماً لترتيب ونشر معلومات عن المحتوى وبنية البيانات الإحصائية.

ISO 11179 - تقنيات المعلوماتية - سجلات البيانات الوصفية (RM)

وضع المعيار العالمي ISO/IEC 11179-1 من طرف اللجنة التقنية JTC1 في الآيزو، (تقنيات المعلومات)، اللجنة الفرعية SC 32 (قسم إدارة البيانات).

”يحدد معيار ISO/IEC 11179 معايير عناصر البيانات وتسجيلها لغاية تسهيل الفهم والتبادل. وتسمح معايير عناصر البيانات وتسجيلها كما هي محددة وفقاً للمعيار ISO/IEC 11179 بإنشاء مجال من البيانات التي يمكن مشاركتها بطريقة أسهل وأسرع بكثير من طرق إدارة البيانات التقليدية“ [9]

وقد أسست بعض المنظمات مفهوم قواعد بياناتها على تعريفات معيار ISO 11179. ولكن ينبغي التشديد على أن هذا الأخير لا يوفر أي أداة عملية للتوثيق والنشر، على عكس المعايير التالية الذكر، المستندي إلى لغة XML.

الإطار 3 إلى من تتوجه مبادرة DDI؟

- مكتب الإحصاءات الوطني، مكتب إحصاءات العمل والتوظيف (Statistics Labour and Employment) في الفلبين (www.bles.dole.gov.ph, www.census.gov.ph, www.bas.gov)
- قسم المسوح والاستقصاءات في سريلنكا (http://statistics.sltidc.lk/)
- وكالة الإحصاءات المركزية في أثيوبيا (www.csa.gov.et)
- وغيرها من مؤسسات كثيرة... (أنظر www.ihsn.org/adp)

المنظمات الدولية:

- اليونسيف، لمنهجية الاستقصاء بمؤشرات متعددة (MICS) (http://www.childinfo.org/mics3_surveys.html)
- البنك الدولي (http://data.worldbank.org/)
- الصندوق العالمي (http://www.theglobafund.org/html/5YEdata/) Global Fund

أصبح اعتماد معيار DDI للبيانات الوصفية أسهل بكثير بفضل البرمجيات السهلة الاستخدام، مثل Metadata Editor DDI وغيره من أدوات الفهرسة المتوافقة مع معيار DDI التي تقدّمها الشبكة الدولية لمسوح الأسر IHSN (أنظر www.ihsn.org/toolkit et www.ihsn)

تستخدم مبادرة DDI للبيانات الوصفية جماعة واسعة من الأرشيفات، لا سيما تلك الخاصة بمكتبات البيانات الجامعية، وإدارات بيانات مكتب الإحصاءات الوطني وغيره من منتجي البيانات الرسميين، بالإضافة إلى المنظمات الدولية.

أمثلة على المستخدمين في الوسط الجامعي:

- DataFirst، جامعة الكاب (www.datafirst.uct.ac.za)
- UKDA، جامعة Essex (www.data.archive.ac.uk/)
- ICPSR، جامعة ميتشيغان (www.icpsr.umich.edu)
- الجامعات الكندية المشاركة في برنامج مركز بيانات البحوث التابع لمركز Statistique Canada (http://www.statcan.gc.ca/rdc-cdr/network-reseau-eng.htm)
- شبكة DataVerse التابعة لمركز بيانات Harvard-MIT ومكتبة جامعة هارفرد (http://thedata.org/)
- الوكالات المنضمة إلى CESSDA (http://www.cessda.org)

منتجو البيانات الرسمية في أكثر من 05 بلداً، نذكر منها:

- Statistique Canada، عن طريق مبادرة إتاحة البيانات (DDI) (http://www.statcan.gc.ca/dli-ild/dli-DDI-fra)
- معهد بوليفيا الوطني للإحصاءات (http://www.ine.gov.bo/anda/)

ومسوح وملفات إدارية وتجارب وملاحظات مباشرة وغيرها من الطرق

مبادرة توثيق البيانات (DDI)

القسم 5.0: وثائق أخرى

يسمح هذا القسم بإضافة وصف ملفاً أخرى على علاقة بالدراسة، مثل: المصادر الوثائقية (نماذج استطلاعات الرأي، ملاحظات الترميز، التقارير التقنية والتحليلية، دليل الشخص الذي يجري المقابلة، إلخ)، وبرامج معالجة البيانات وتحليلها، والصور والخرائط.

معيار دبلن كور *Dublin Core* للبيانات الوصفية (DCMI)

اقتطف محتوى هذا القسم من الموقع الإلكتروني الخاص بمبادرة DCMI (DCMI-ive) <http://dublincore.org>: (Dublin Core Metadata Initiat (متوفر بالإنجليزية فقط)

يسمح مجموع عناصر البيانات الوصفية الخاصة بـ *Dublin Core* (معيار ISO 15836)، المسمى أيضاً معيار بيانات *Dublin Core* الوصفية، بوصف المصادر الإلكترونية. وتتجلى أهمية هذا المعيار في وصف المصادر الكامنة في البيانات الوصفية: نماذج استطلاعات الرأي والتقارير والإرشادات والنصوص وبرامج معالجة البيانات، إلخ. وقد تبلورت فكرة وضع هذا المعيار عام 1995، برعاية مركز OCLC (Online Computer Library Center) ومركز NCSA (National Center for Supercomputing Applications) اللذين اجتمعا في دبلن (أوهايو). وتدرجياً، أصبح *Dublin Core* المعيار الأكثر انتشاراً لوصف المصادر الرقمية على الشبكة، وقد رفع مستواه إلى مستوى معيار ISO عام 2003، وتضمن مبادرة DCMI الحفاظ على المعيار ويعمل على تطويره. كما تهدف هذه المنظمة العالمية إلى الترويج لوضع معايير البيانات الوصفية القابلة للتشغيل المتبادل.

يعود الفضل في نجاح معيار *Dublin Core* للبيانات الوصفية إلى بساطته في المقام الأول. فمنذ البداية، حرص مصمموه على أن يبقى مجموع العناصر مختصراً وبسيطاً بقدر الإمكان، حتى يتمكن الأشخاص من غير ذوي الخبرة من استخدامه، ويهدف المعيار إلى تشجيع إنشاء ملاحظات وصفية بسيطة بسهولة وبتكلفة منخفضة للمصادر التربوية، وإلى إتاحة البحث عن هذه المصادر نفسها بفعالية على شبكة الإنترنت أو في أي مجال على شبكة. ويحتوي *Dublin Core* بنسخته الأبسط على 15 عنصراً من البيانات الوصفية، كلها تشغيلية وتستخدم عدة مرات. أما هذه العناصر فهي:

- | | | |
|---------|----------|---------|
| العنوان | العلاقة | الحقوق |
| الموضوع | التغطية | التاريخ |
| الوصف | المنشئ | الصيغة |
| النوع | المراجع | المعرّف |
| المصدر | المتعاون | اللغة |

الإطار 4 إلى من يتوجه؟ *Dublin Core*

Dublin Core هو معيار بيانات وصفية فائق في السلاسة وسهولة الاستخدام. وهو ليس مفضلاً بقدر معيار DDI (إذ لا يحتوي *Dublin Core* على عناصر تتعلق بمعلومات عن ملفات البيانات والمختبرات)، غير أنه يمكن أن يصف السمات العامة لمجموع البيانات. كما أنه يوفر خياراً يسمح بوصف مجموعات بيانات لا يمكن لمصادرها أن تخضع لتوثيق مفصل. هذا وتستخدم نسخة معدلة من *Dublin Core* في open data initiative في الولايات المتحدة (Data.gov). (www)

المنهجية المطبقة للحصول على قياسات تجريبية. وبعبارة أخرى، يمكن لوحة التحليل أن تكون مؤلفة من أفراد، أو أسر، أو عائلات، أو شركات، أو صفقات، أو بلدان أو حتى مواضيع أخرى تمثل اهتماماً علمياً. ومن ناحية أخرى، يمكن أن تتمثل الملاحظات بقياسات تجري بدقة في مكان معين (خلال أسبوع ضمن عينة تمثل بلداً على سبيل المثال). كما يمكن أن تكون الملاحظات متكررة ومعدة في أماكن مختلفة (بيانات طولية ومستعرضة من بلدان مختلفة، وبيانات متسلسلة زمنياً ومعتمدة). وتحتوي مبادرة DDI على تعريف مسهب منهجية الدراسة (طريقة جمع البيانات، وطرق أخذ العينات عند الاقتضاء، وتحديد المكان، والمناطق الجغرافية المدروسة، والجهة والأشخاص المسؤولين، إلخ).

الهيكلية

تسمح مبادرة DDI بوصف كل نواحي الاستقصاء بالتفصيل: من المنهجية إلى المسؤوليات والملفات والمختبرات. وهي توفر قائمة مهيكلة ومسهب من بضع مئات العناصر والخواص القادرة على تشكيل التوثيق لمجموع البيانات. وإذا كان من غير المرجح أن تظهر كلها، فإن ظهور بعض العناصر مثل العنوان (Title) إلزامي، وأحادي بالضرورة. أما العناصر الأخرى مثل الكاتب والباحث الرئيسي (Primary Investigator) فهي اختيارية ويمكن أن ترد أكثر من مرة، لأنها تتعلق بالشخص / الأشخاص و/أو المؤسسة / المؤسسات المسؤولة عن الاستقصاء.

تبيّن عناصر معيار DDI (نسخة 2.4) في خمسة أقسام:

القسم 1.0: وصف الوثيقة

ليست الوكالة التي تتكلف بتوثيق الدراسة ونشرها (أو الاستقصاء أو المسح) هي نفسها التي تنتج البيانات دائماً. لذا، من الضروري تقديم معلومات (بيانات وصفية) عن الدراسة بحد ذاتها، وكذلك عن عملية التوثيق. ويجب أن يحتوي قسم "الوصف" (Description) على تعريف عام يصف ملف XML بصيغة DDI، أي أن يقدم "بيانات وصفية عن بيانات وصفية أخرى".

القسم 2.0: وصف الدراسة

يقدم هذا القسم تعريفاً عاماً بالدراسة، وهو يحدد مراجع الاستقصاء وهوية الشخص الذي تلقى البيانات أو جمعها، وهوية الشخص الذي نشرها. كما يلخص هذا القسم محتوى البيانات، ويعطي معلومات عن طرق جمعها ومعالجتها، إلخ.

القسم 3.0: وصف ملفات البيانات

يصف هذا القسم محتوى كل ملف من البيانات، ويحدد عدد التسجيلات والمختبرات، كما يذكر النسخة، واسم المنتج، إلخ.

القسم 4.0: وصف المختبرات

يتعلق هذا القسم بمعلومات كل من المختبرات، لا سيما بصياغة الأسئلة، وبيئة الاستقصاء، وعناوين المختبرات وقيمها، وطرق الاشتقاق والإلصاق.

الاطاء، 5 لغة XML

```
<collDate date="2005-01" event="start"/>
<collDate date="2005-03" event="end"/>
<nation>Popstan</nation>
</geogCover>geogCover</geogCover>
</sampProc>5000 أسرة؛ على مستويين</sampProc>
</respRate>98 بالمئة</respRate>
```

إن استخدام البطاقات أكثر فعالية حين تعتمد جماعة مجموعة مشتركة من البطاقات (معايير DDI أو Dublin Core مثلاً). يقدّم اعتماد مجموعة مشتركة من بطاقات XML مزايا كبيرة لتوثيق البيانات الجزئية:

- إنشاء قائمة شاملة بعناصر البيانات الوصفية المفيدة؛
- إمكانية تقييم محتوى ملف مع التأكد مما إذا كان يحتوي على بطاقات بيانات أو لا؛
- إنشاء فهرس لمجموعات البيانات يسمح بالبحث عن عناصر البيانات الوصفية الرئيسية؛
- إمكانية تحويل الملف إلى صيغ أسهل استخداماً.

يمكن تحويل ملفات XML إلى صيغة HTML، أو PDF أو إلى أنواع أخرى من المستندات التي تنفذ تحويلات LSX. ويمكن عندئذٍ تبادلها من شبكة إلى أخرى أو على الإنترنت عن طريق خدمات beW أو عن طريق بروتوكول PAOS (وهو بروتوكول مبني على لغة XML يسمح بتبادل المعلومات بين التطبيقين عن طريق HTTP). أما صفحة HTML أدناه فهي مثال عن تحويل XSL للنسخة الإنجليزية من ملف XML أعلاه:

يمثل لفظ XML اختصاراً لـ eXtensible Markup Language. وتقدّم هذه اللغة سبباً لوضع بطاقة على نصوص بواسطة بطاقات تحدد معنى هذه النصوص ولا تكتفي بالتعريف بها. وبعبارة أخرى، تسمح صيغة XML بهيكلية نص بواسطة بطاقات تحتوي على مؤشرات دلالية. وبالمبدأ، فيما يتعلق بالتنظيم، توازي هذه "البطاقات" "الحقول" في قاعدة بيانات. ولكن على عكس ملفات قاعدة البيانات، يمكن لملفات XML أن تعرض وتعدّل بواسطة مراجع نصوص معياري. ويمكن أن يجرى البحث في هذه الملفات بنفس الطريقة التي يجرى بها في قاعدة بيانات تقليدية، بواسطة الأدوات الملائمة. وكما يمكن أن يسمح محتوى قاعدة بيانات بإنشاء تقرير، يمكن أن تعرض ملفات XML وتحوّل بمساعدة تطبيقات برمجيات أخرى للحصول على صيغ أسهل استخداماً (مثل صحائف الجدولة أو صيغة PDF أو صفحات الويب).

يسمح المثال أدناه بتكوين فكرة عن الطريقة التي يمكن فيها عرض المعلومات النصية عن الاستقصاء بصيغة XML:

المعطيات: من يناير/ كانون الثاني حتى مارس/ آذار 2005، أجرى مكتب الإحصاءات الوطني بوبستان Popstan استقصاءً متعدد المؤشرات (MCIS) ممولاً من اليونيسيف. وتم اختيار 5000 أسرة تمثل مجموع سكان البلد عشوائياً للمشاركة في الاستقصاء، وفقاً لمخطط أخذ عينات ذي مستويين. وقد قدمت 4900 أسرة منها معلومات.

وتظهر المعلومات نفسها محولة إلى صيغة DDI-XML على الشكل التالي:

```
<titl>استقصاء متعدد المؤشرات 2005</titl>
<altTitl>MICS</altTitl>
</AuthEnty>مكتب الاحصاءات الوطني</AuthEnty>
</fundAg>fundAg abbr="منظمة الأمم المتحدة للطفولة"
```

3. ما هي الحجج التي تدعم بث البيانات الجزئية؟

”تمثل مهمة الأرشيف الوطني للبيانات الجزئية في سريلنكا بالاستجابة إلى حاجات جماعة الباحثين الوطنية والدولية، التي تجهد بحثاً عن إجابات على المشاكل الاجتماعية الاقتصادية التي يواجهها العالم. وبمساعدة قسم التعداد والإحصاءات السريلنكي Department of Census and Statistics، تنشر LankaDatta بيانات إحصائية ثابتة وموثوق بها وحديثة، تنتجها وكالات النظام الإحصائي الوطني، كما تحفظ سرية هوية المستجيبين تماماً“ . Department of Census and Statistics، سريلنكا / [http://statistics](http://statistics.sltidc.lk)

يمثل البث أحد المسؤوليات الأساسية التي يقوم بها معهد إحصاءات. ويعد هذا الفصل الإيجابيات الكثيرة التي يوفرها نشر البيانات الجزئية. إذ تظهر دراسة بيانات خاصة بهذه المهمة، وسياسات البث، إلى جانب تجارب منتجي بيانات مختلفين حول العالم أهمية إتاحة إمكانية النفاذ إلى ملفات البيانات الجزئية والحجج التي تدعم هذا الموقف.

3.1.1 دعم البحث

”تشكل مبادرة إتاحة البيانات (DDI) مثلاً على إعطاء التكنولوجيا حقها، تكنولوجيا سير المعلومات في كندا، إذ أنها تسمح للمؤسسات التعليمية للمرحلة ما بعد الثانوية بأن تقدّم لأول مرة باقة كاملة من خدمات البيانات إلى الطلاب والأساتذة على حد سواء. ويبدو جلياً أكثر فأكثر أن المبادرة تساهم بشكل كبير في التعليم والبحث في كندا. (...) هذا وقد ساهمت مبادرة إتاحة البيانات في ظهور ثقافة البيانات في كندا“ . (Statistique Canada <http://www.statcan.gc.ca/dli-ild/about-apropos-fra.htm>) ؛ أنظر أيضاً Watkins (31)

لعل السبب الرئيسي (وغالباً ما يكون الدافع الصريح الأوحى) الذي يدفع منتجي البيانات إلى مشاركة بياناتهم الجزئية هو دعم الأبحاث. ففي نهاية الاستقصاء، عادة ما ينتج المسؤولون عن جمع البيانات سلسلة من الجداول المعدة لإظهار الجوانب الأساسية وتقديم لمحة عامة عن النتائج إلى المستخدمين. ولا يمكن أن نتظر من هذه الوكالات أن تحدد كل مواضيع الأبحاث التي يمكن لهذه البيانات أن تساعد في معالجتها، إذ لا تملك الميزانية اللازمة لذلك. وتوفر ملفات البيانات الجزئية للباحثين بعداً مهماً لتحديد العلاقات والتفاعلات بين الظواهر التي ستناولها استقصاء ما، ما يعزز تعددية أعمال الأبحاث وجودتها⁵.

3.2. تعزيز مصداقية الإحصاءات الرسمية

يشهد توسيع نطاق النفاذ إلى البيانات الجزئية على ثقة المنتجين ببياناتهم، إذ يسمحون بإصدارها مجدداً أو تصديقها من جهة منظمات مستقلة.

هنا بعض الأمثلة التي تظهر كيف تصف بعض معاهد الإحصاءات الوطنية سياسة بث البيانات التي تعتمد عليها.

3.3. تحسين موثوقية البيانات وثباتها

تتمثل فائدة أخرى في العلاقة الأقرب بين منتجي البيانات والمستخدمين المرشدين. فغالباً ما يكون استخدام البيانات العامل المساهم في تكوين فكرة عن التحسينات الممكنة، لا سيما على صعيد صياغة الاستقصاءات ونشر البيانات الجزئية. ويمكن لعملية نشر البيانات الجزئية أن تشمل عملية نمطية لحصد آراء الباحثين، كما هو الحال في Census Bureau الأمريكي. فقد تؤدي آراء المستخدمين إلى تحسين تدريجي للاستقصاءات.

”يدعم النفاذ إلى البيانات الجزئية ويشجع اتخاذ قرار يسمح باستخدام واسع لبيانات مكتب الإحصاءات الأسترالي (ABS Bureau of Statistics) لغايات تحليلها والبحث الاقتصادي والاجتماعي. فمذ العام 1985، يتيح مكتب الإحصاءات الأسترالي بيانات جزئية بيانات جزئية بشكل ملفات تسجيلات وحدوية سرية (CURF)، بشروط معينة ولغايات إحصائية. واليوم، أصبح مكتب الإحصاءات الأسترالي يتلقى طلبات كثيرة للنفاذ إلى تسجيلات وحدوية أكثر تفصيلاً، بأشكال أكثر سلاسة تصدر عن نطاق أوسع من مجموعة واسعة من البيانات (بيانات شركات وبيانات طويلة على سبيل المثال). وإن أي استحالة لتلبية الطلبات سيكون لها وقع سلبي متنام على الأنشطة الرئيسية لمكتب الإحصاءات الأسترالي، وعلى مصالحه، وأخيراً على تناسق نظام الإحصاءات الوطني. وقد دفع هذا التطور المقترن بسلسلة من العوامل الأخرى التي تبرر التغيير، لا سيما الخطر المتنامي من حيث التعرف على هوية المستجيبين، مكتب الإحصاءات الأسترالي إلى اقتراح استراتيجية جديدة للنفاذ إلى البيانات الجزئية في المستقبل“ . (مكتب الإحصاءات الأسترالي، <http://www.abs.gov.au>، ترجمة من الإنجليزية).

3.4. تقليص تكرار البيانات

غالباً ما يثني وضع ملفات البيانات الجزئية في تصرف المستخدمين هؤلاء الأخيرين عن جمع البيانات التي يحتاجون إليها بأنفسهم. وهكذا تقل الطلبات المتوجهة إلى المستجيبين ويتقلص خطر الحصول على دراسات غير متناسقة حول الموضوع عينه.

3.5. تحسين مردود الاستثمار

”لا يساهم نفاذ العامة إلى بيانات الأبحاث الممولة من الأموال العامة ومشاركتها في تعزيز تأثير التكنولوجيا الحديثة والشبكات الرقمية الجديدة على قدرة البحث فحسب، بل يسمحان أيضاً بتحسين مردود الاستثمار العام في الأبحاث أيضاً...“

”يتمثل الهدف الأساسي لمكتب الإحصاءات المركزية (CSO Office Central Statistics) في ما يتعلق بالنفاذ إلى البيانات الجزئية في دعم جماعة الباحثين والتأكد من أن البيانات التي نجمعها مفيدة لأقصى حد ممكن. وتعدّ هذه المقاربة أساسية لوضع سياسات مبنية على عناصر مقنعة. كما يمكن أن تساهم في تقليص تكاليف البحث وتفاذي تكرار البيانات“ . (المكتب المركزي للإحصاءات الإيرلندية، <http://www.cso.ie>، ترجمة من الإنجليزية).

إذ يجمع الباحثون ومؤسسات البحث التي تمويلها الأموال العامة كميات متنامية من البيانات. وتمثل كميات بيانات الأبحاث هذه التي تتوسع بسرعة استثماراً هائلاً للأموال العامة وأحد الموارد الممكنة للمعارف اللازمة لمعالجة تحديات عديدة تواجه الإنسانية...“

5 يجسد عمل هاميلتون وهامفري جيداً الاستخدام العميق الذي يمكن أن يجريه الباحثون لملفات البيانات الجزئية. وهو يبيّن أن يضع مئات مشاريع الأبحاث أجريت على قاعدة الاستقصاء الوطني بشأن صحة السكان (ENSP) في كندا، بعد نشره بشكل ملف بيانات جزئية متاحة للاستخدام العام. أنظر أيضاً [30]

ومن أجل تعزيز مردود علمي واجتماعي أفضل للاستثمارات العامة في بيانات الأبحاث، وضعت الدول المتنامية إلى منظمة التعاون والتنمية الاقتصاديين، على سبيل المثال، عدة قوانين وسياسات وممارسات بشأن النفاذ إلى بيانات البحث على المستوى الوطني. وفي هذا السياق، تشكل الخطوط التوجيهية الدولية ورقة رابحة لتشجيع تبادل بيانات الأبحاث واستخدامها على المستوى العالمي” [17]

غالباً ما تسمى تطبيقات ويب هذه المبتكرة التي تستخدم بيانات مفتوحة باسم mashups (أو التطبيقات المركبة).

الإطار 6 الالتزام القانوني بثّ البيانات الجزئية: مثال المركز الوطني للإحصاءات الصحية (الولايات المتحدة)

يعطي المركز الوطني للإحصاءات الصحية الأمريكي (NCHS)، المرفق بمركز الحد من الأمراض والوقاية منها (for Disease Control and Prevention US Centers) مثلاً عن تشريع يحكم نشر البيانات الجزئية.

”ويتعيّن على المركز الوطني للإحصاءات الصحية الأمريكي، بصفته المعهد الفدرالي للإحصاء بذل كل ما في وسعه لتحقيق أقصى قدر من توافر البيانات، بما في ذلك تقصير الفترة الفاصلة بين جمع البيانات ونشرها، وذلك في سبيل تحسين الجودة، وتقليل مخاطر إفشاء هوية المستجيبين في آن معاً. كما تنصّ التشريعات التي تنظم أنشطة المركز الوطني للإحصاءات الصحية الأمريكي على إلزامية وضع البيانات في تصرف المستخدمين على أوسع نطاق ممكن (...).

”ومع ذلك، ينبغي النظر إلى الالتزام بنشر البيانات من منطلق جودة المعهد الفدرالي للإحصاء للمركز الوطني للإحصاءات الصحية الأمريكي، كما يجب أن تتماشى مع الحاجة إلى حماية سرية المستجيبين وضمان جودة البيانات. (...)

”إن القانون الذي يلزم المركز الوطني للإحصاءات الصحية الأمريكي بنشر بياناته هو نفسه الذي يجبره على الحفاظ على سرية هوية الأفراد أو المنظمات التي تظهر في أنظمة البيانات الخاصة به“. [14] (ترجمة من الإنجليزية)

في مجال تطوير المواقع الإلكترونية، يتمثل التطبيق المركب mashup في صفحة ويب أو تطبيق يستخدم أو يمزج بيانات أو وظائف من مصادر خارجية عدة بهدف إنشاء خدمة جديدة. ويستدعي ذلك تكاملاً سهلاً وسريعاً، غالباً ما يتم عن طريق واجهات برمجة التطبيقات، وباستخدام مصادر بيانات مفتوحة. أما الهدف فيمكن في الحصول على نتائج تغني البحث، ولا تنحصر بالضرورة في الدافع الأساسي لإنتاج البيانات البحثية الأولية.

ولتحقيق النفاذ الدائم إلى بيانات الخدمات الأخرى، صممت التطبيقات المركبة mashups لتكون تطبيقات للملّم أو تطبيقات ترعى على شبكة الانترنت. كما تؤدي تطبيقات mashups دوراً فاعلاً لا شك في تطور البرمجيات الاجتماعية وشبكة الويب 2.0 (انظر أيضاً http://fr.wikipedia.org/wiki/application_composite).

يهدف تشجيع تطوير ال mashups لم يكتفِ منتجو البيانات بنشرها، بل أصبحوا يقدمون أيضاً تطبيقات برمجية تدعى واجهات برمجة التطبيقات (API). وتسمح هذه الأخيرة للمطوّرين باستخدام البيانات بسهولة أكبر.

ومن أجل تعزيز مردود علمي واجتماعي أفضل للاستثمارات العامة في بيانات الأبحاث، وضعت الدول المتنامية إلى منظمة التعاون والتنمية الاقتصاديين، على سبيل المثال، عدة قوانين وسياسات وممارسات بشأن النفاذ إلى بيانات البحث على المستوى الوطني. وفي هذا السياق، تشكل الخطوط التوجيهية الدولية ورقة رابحة لتشجيع تبادل بيانات الأبحاث واستخدامها على المستوى العالمي” [17]

3.6 جمع الأموال للدراسات الإحصائية

كلما نشرت ملفات البيانات على نطاق أوسع واستخدمت، زادت قيمتها لدى الممولين.

يمكن لهذه الحجة أن تشجع الرعاة على تمويل جمع البيانات. كما أن بعض الممولين يشترطون دليلاً يثبت استخدام البيانات.

يؤدي الاستخدام الأفضل للبيانات إلى مردود أفضل على الاستثمار بالنسبة إلى الرعاة، الذين سيتشجعون أكثر على دعم أنشطة جمع البيانات. وقد أصبح تمويل المنظمات الدولية للاستقصاءات مشروطاً أكثر فأكثر بنشر مجموع البيانات التي يتم الحصول عليها.

3.7 تقليص تكلفة نشر البيانات

أخيراً، يمكن لوكالات جمع البيانات أن تستفيد من فعالية متنامية، إذ ستتمكن من تقليص كمية الجداول المنتجة المحددة مسبقاً والتركيز أكثر على وضع بيانات التلخيص. إن التشديد على هذا الجانب، لا سيما في وسائل الإعلام وقطاع التعليم، يمكن أن يجذب اهتمام جمهور أكبر ويسهل تمويل أنشطة وكالات جمع البيانات، سواء أكانت مكاتب إحصاءات وطنية أم غيرها من المؤسسات. غير أن هذه الجداول غير كافية للتحاليل المعمقة. وأخيراً، ينبغي التفكير في مسألة الفعالية في سياق تكاليف إنتاج ملفات البيانات الجزئية ونشرها. وستتناول هذه المسائل في الفقرة التالية.

3.8 احترام الإلتزامات القانونية والتعاقدية

في بعض البلدان، تجبر بعض الوكالات العامة على نشر قسم من بياناتها الجزئية. وغالباً ما يمول المكلفون جمع البيانات، لذا يعتبر هذا الأمر جزءاً من القطاع العام. وفي بلدان أخرى، يتمثل الممولون في منظمات راعية، تشترط أن توضع البيانات التي يتم الحصول عليها في تصرف الباحثين (انظر الإطار 6). ولا يتعارض هذا الالتزام بنشر البيانات الجزئية مع الالتزام باحترام سرية الغير وحياتهم الخاصة. بل يتعين على كل مسؤول عن الإحصاءات في كل مؤسسة، أو في لجنة إفشاء السرية، أن يقرر محتوى البيانات الجزئية التي تنشر، بالإضافة إلى الإجراءات التي يجب وضعها لإنشاء ملفات متاحة للاستخدام العام.

3.9 تشجيع تطوير أدوات جديدة لاستخدام البيانات

تنتشر صيحة جديدة منذ بضع سنوات، هي صيحة البيانات المفتوحة. أما المفهوم الرئيسي وراء هذه البيانات هو أن جمعها يحوّل عن طريق الأموال العامة أو تحت إشراف وكالة للدولة تعنى بالقطاع العام.

وقد تبنت حكومات عديدة هذا الاتجاه، مثل nment Initiative - Open Gover في الولايات المتحدة (www.data.gov) أو مثيله البريطاني

برمجة التطبيقات. http://fr.wikipedia.org/wiki/Interface_de_programmation.
تمت زيارة الموقع في 9 يناير / كانون الثاني 2011:
أنظر أيضاً الإطار (7)

”تمثل واجهة برمجة التطبيقات (API) و Programming Interface واجهة تقدّمها برمجة معلوماتية. وهي تسمح بتفاعل البرامج بعضها مع بعض، بحيث تشبه واجهة إنسان مقابل الحاسوب، ما يجعل التفاعل ممكناً بينهما“. وقد بدأ العمل بواجهات برمجة التطبيقات (أو وضعها) عن طريق تطبيق، أو مكتبة برمجيات أو نظام تشغيل. وهي تتيح تحديد تركيب الطلب واتفاقياته التي يتعين على المبرمج احترامها لاستخدام الخدمات الموضوعية في تصرفه. كما يمكن أن تضم خاصيات تتعلق بروتين التنفيذ، وهيكل البيانات، وفئات المواضيع، وبروتوكولات التواصل بين المستخدم والمركّب الذي يضع واجهة

الإطار 7 الترويج لتطوير التطبيقات المركبة spuhsam عن طريق بث البيانات المفتوحة وواجهات برمجة التطبيقات

نعلم أن أشخاصاً كثيرين مثلكم، من خارج الحكومات، يملكون المهارات والقدرات اللازمة للاستفادة القصوى من البيانات العامة. ويشكل هذا بالنسبة إلينا نقطة الإنطلاق لعلاقة التعاون التي تأمل تكوينها معكم.
المصدر: <http://data.gov.uk>. تمت زيارة الموقع في 3 أبريل / نيسان 2010، ترجمة من الإنجليزية

تتناول مبادرة البيانات المفتوحة موضوع النفاذ. مع إطلاق donnees.banquemondiale.org، أصبحنا نكثف جهود البنك الدولي في سبيل فتح فهارس بياناته عن طريق نفاذ سهل للغاية ومباشر على شبكة Web. وسنقوم خلال عام 2010 بنشر موجتين من ميزات Web ستمثلان منصة لهذا الجهد. (...)

أما المرحلة الثانية فستتمحور حول الترويج لاستخدام بيانات البنك الدولي بواسطة واجهة برمجة التطبيقات على شبكة الويب. وسيستدعي ذلك إجراء تحسينات على واجهة برمجة التطبيقات الحالية، بالإضافة إلى إتاحة مجال للتواصل حول موضوع واجهة برمجة التطبيقات و عتبة أدنى للدخول، لا سيما بالنسبة إلى المستخدمين من غير مطوّري المعلوماتية.

وستعلن هذه المرحلة عن بداية منعطف يحوّل واجهة برمجة التطبيقات على شبكة الويب من واجهة موجهة حصراً للمطوري المعلوماتية إلى واجهة تتوجه أيضاً إلى الباحثين وصانعي القرارات السياسية إلى جانب مطوّري المعلوماتية.
المصدر: <http://donnees.banquemondiale.org/developpeurs>. تمت زيارة الموقع في 7 مايو/ أيار 2010

الولايات المتحدة: data.gov - اكتشاف، مشاركة، التزام

يتمثل أحد الأهداف الكامنة بمبادرة Open Government Initiative في تعديل عادات نشر المعلومات عن طريق إضفاء طابع مؤسسي على إتاحة البيانات الفدرالية على نطاق واسع، وذلك في صيغ تسهل النفاذ إليها. يمثل موقع data.gov المنتج الريادي لمبادرة Open Government Initiative، وهو يهدف إلى تسهيل النفاذ إلى مجموع البيانات الفدرالية التي تسمح للجمهور الكبير بفهم طريقة عمل الوكالات الفدرالية وأنشطتها بشكل أفضل، وتسلط الضوء على مهماته، وتدير الفرص الاقتصادية، وتحسن شفافية الحكومة الفدرالية وانفتاحها (مجموع البيانات ذات القيمة المضافة العالية).

هينياً للمطوّرين

نحن مدهولون أمام العدد المسجل لعمليات التحميل وللتطبيقات المبتكرة التي تحتوي على بيانات حكومية ظهرت منذ إطلاق [Data.gov](http://data.gov). ولعل أحد الأهداف الأساسية لموقع [Data.gov](http://data.gov) يتمثل في تشجيع الابتكار. وفي هذا الصدد، تميّزت جماعة مطوّرينا أيما تميّز. إننا نشجع كل المطوّرين والمبرمجين على اكتشاف مجموع البيانات المدرجة في [Data.gov](http://data.gov) وعلى المشاركة الفاعلة في هذه الجماعة الحيوية والمتنامية.

المصدر: <http://www.data.gov/open>. تمت زيارة الموقع في 3 أبريل / نيسان 2010، ترجمة من الإنجليزية

4. ما هي التكاليف والمخاطر المرتبطة بثّ البيانات الجزئية وكيف يمكن ضبطها؟

يتعيّن على مكاتب الإحصاء الوطنية وغيرها من منتجي البيانات / جامعيها أن تأخذ في الاعتبار عدداً من العوامل لتحديد وتنفيذ سياساتها وبرامجها الرامية إلى بثّ البيانات الجزئية؛ ويتضمن ذلك التكاليف والمهارات، والمسائل المتعلقة بجودة البيانات، وإمكانية إساءة استغلال المستخدمين للبيانات أو إساءة تحليلها، والنواحي القانونية والأخلاقية، ناهيك عن الحفاظ على ثقة المستجيبين ودعمهم.

4.1.1 المسائل الأخلاقية والحفاظ على سرية المستجيبين

عند جمع المعلومات من الأفراد والشركات أو المنظمات الأخرى، عادة ما تضمن معاهد الإحصاءات وغيرها من منتجي البيانات للمستخدمين أن المعلومات المقدمة ستستخدم لأغراض إحصائية فقط. وهنا التزام معنوي وأخلاقي.

”يجب أن تحافظ مكاتب الإحصاء الوطنية على ثقة المستجيبين حتى يستمر هؤلاء بالتعاون في جمع البيانات الخاصة بهم، وتشكل حماية السرية الجانب الرئيسي من هذه الثقة. فإذا اعتقد المستجيبون أو شعروا بأن مكاتب الإحصاء الوطنية ليست قادرة على ضمان سرية البيانات، من غير المرجح أن يتعاونوا أو يقدموا معلومات دقيقة. فحتى الحوادث المعزولة، حين تكون مصحوبة بتغطية إعلامية قوية، يمكن أن يكون لها تأثير كبير على تعاون الأشخاص المستجيبين، وبالتالي على نوعية الإحصاءات الرسمية، تمثل حماية السرية الشاغل الرئيسي لمكاتب الإحصاء الوطنية، ولكنها ليست الوحيدة. إذ تتمتع هذه المكاتب بالسلطة الكافية، سواء عن طريق تفويض قانوني أو أي شكل آخر من أشكال الترخيص، لدعم نفاذ الباحثين إلى البيانات الجزئية.“ [24] (الترجمة من الإنجليزية)

لتحقيق الاستفادة القصوى من البيانات الجزئية، يتعيّن على مكاتب الإحصاءات الوطنية وغيرها من منتجي البيانات إيجاد توازن بين تلبية متطلبات السرية وتوفير حقوق النفاذ، ولعلّ أحد الحلول يتمثل في استخدام أنواع مختلفة من البيانات الجزئية (أنظر الفصول السابقة). وثمة خيار آخر غير ممكن في جميع الحالات؛ وهو أن يعطي كل مستجيب موافقته الرسمية على نشر البيانات المجموعة.

الحصول على موافقة الأشخاص الماديين

”يمكن الحصول على الموافقة عن طريق التوقيع أو ضمناً، في الحالة الأولى، يتم إبلاغ المستجيب خطأً بشروط الاستخدام المحددة مسبقاً للمعلومات المطلوبة منه، وهو مدعو إلى الموافقة عن طريق توقيع الوثيقة. ومع ذلك، لا يتم إبلاغ المستجيب دائماً خطأً، بل يتم أحياناً شفهاً فحسب. وإذا وافق على تقديم المعلومات المطلوبة، ”يستنتج“ [الشخص الذي يجري الاستقصاء] أنه يوافق على الاستخدام المقترح للبيانات وإيصالها إلى الأطراف المحددة خطأً أو المذكورة شفهاً. ويكون [للشخص الذي يجري الاستقصاء] الحق في استخدام البيانات بالطريقة التي حددت للمستجيب حصراً.“ [15]

الحصول على موافقة الأشخاص المعنويين

”في حالة الأشخاص المعنويين، تختلف الطريقة إذا كانت المعلومات مطلوبة شفهاً أو عن طريق البريد الإلكتروني.

أ. إذا كان عضو من فريق العمل أو وكيل للجهة التي تجري الاستقصاء يطلب المعلومات شفهاً، يتعيّن عليه أولاً أن يحدد الشخص المخول بتقديم المعلومات باسم المؤسسة. عندما يجب هذا الشخص على الأسئلة بعد أن يتم إبلاغه بالاستخدام المرجو للبيانات المجموعة، يعتبر ممثل [وكالة الاستقصاء] أن المؤسسة توافق على هذا الاستخدام.

ب. بالنسبة إلى البيانات التي تجمع عن طريق البريد الإلكتروني، يمكن توجيه الطلب إلى المؤسسة، أو مديرها أو أي شخص آخر تحدده مسبقاً [وكالة الاستقصاء] على أنه مخول بتقديم المعلومات المطلوبة. وتحدد الرسالة التي تحتوي على طلب الاستخدام المرجو للمعلومات. عند تلقي البيانات، يعتبر طاقم عمل [وكالة المقابلة] أن المؤسسة تقبل استخدامها وفقاً للطريقة التي أبلغت بها.“ [15]

4.2 النواحي القانونية

هل يحق لمنهج البيانات من الناحية القانونية نشر ملفات البيانات الجزئية؟ لا يوجد جواب واحد على هذا السؤال. فالتشريعات التي تنظم أنشطة منتجي البيانات خاصة بكل بلد وكل برنامج إطار (الإطار 8).

كما ذكرنا سابقاً، يمثل نشر البيانات الجزئية في بعض الأحيان التزاماً قانونياً. وفي معظم الحالات، تفرض التشريعات قيوداً بسيطة. وبالتالي، تحدد سياسة البلد في مجال نشر البيانات الجزئية وفقاً لإطارها التشريعي. فمن الضروري لمنتجي البيانات أن ”يتمتعوا بأساس قانوني وأخلاقي متين (فضلاً عن الأدوات التقنية والمنهجية اللازمة)، وذلك لحماية السرية. يتطلب هذا الأساس القانوني والأخلاقي أن نوازن بعناية بين المصلحة العامة المتمثلة بحماية فائقة للسرية من جهة، وبين الفوائد التي يوفرها البحث، من جهة أخرى. أما القرار الذي يقضي بالسماح للباحث بالنفاذ إلى البيانات أو عدم السماح له بذلك، فقد يرتبط بمزايا مشروع بحثه ومصداقيته. كما ينبغي أن يؤخذ ذلك بالاعتبار بطريقة أو بأخرى في الأحكام التشريعية.“ [24]

”ينبغي أن تحترم أحكام النفاذ إلى البيانات الحقوق والمصالح المشروعة لجميع أطراف نشاط البحوث العامة. كما سيكون النفاذ إلى بعض بيانات البحث واستخدامها محدودين حتماً نظراً إلى أنواع مختلفة من المتطلبات القانونية، والتي قد تفرض قيوداً للأسباب التالية:

- الأمن القومي: يمكن أن تصنف بعض البيانات المتعلقة بالاستخبارات، أو الأنشطة العسكرية أو صنع القرار السياسي على أنها سرية، وبالتالي يكون النفاذ إليها محدوداً.
- حماية البيانات الخاصة والسرية: إن إمكانية النفاذ محدودة إلى البيانات المتعلقة بالمواضيع الإنسانية وغيرها من البيانات الشخصية وذلك بموجب القوانين والسياسات الوطنية التي ترعى حماية السرية والحياة الخاصة. ولكن يتعيّن على أصحاب هذه البيانات أن يضعوا أحكاماً تحفظ السرية ومجهولية الهوية لضمات مستوى مرض من السرية من أجل الحفاظ على نفعية البيانات بالنسبة إلى الباحثين بأكثر قدر ممكن.
- الأسرار التجارية وحقوق الملكية الفكرية: يمكن أن يتعدّد نفاذ الباحثين إلى البيانات المتعلقة بشركات أو أطراف أخرى، أو صادرة

إذا لم يكن الاستخدام المرجو للبيانات الجزئية متماشياً مع أغراض إحصائية أو تحليلية، لا ينبغي منح حق النفاذ إلى البيانات الجزئية. ويمكن للجان الأخلاقيات، أو أي نظام مماثل لها، أن تقدم المساعدة في حال الشك بشأن النفاذ إلى البيانات الجزئية.

يستخدم الباحثون البيانات الجزئية لإجراء البحوث. ويشمل هذا تطوير المجاميع الإحصائية مختلف أنواعها، أو إنشاء التوزيعات الإحصائية، أو تعديل النماذج الإحصائية، أو تحليل إحصاءات مختلفة تعود إلى مجموعات سكانية فرعية، وتتماشى هذه الاستخدامات مع الأهداف الإحصائية. وفي هذا الصدد، يمكن اعتبار البيانات الجزئية مفيدة لدعم البحوث^[24].

المبدأ الثالث: على تقديم البيانات الجزئية أن يتوافق مع الأحكام القانونية وغيرها من الأحكام التي تضمن سرية البيانات الجزئية المقدمّة

”فيما يتعلق بالمبدأ الثالث، تقضي الحالة المثالية بوضع أحكام قانونية تهدف إلى حماية السرية قبل نشر أي بيانات جزئية كانت، ومع ذلك، يجب أن تستكمل الأحكام القانونية بتدابير إدارية وتقنية لتنظيم النفاذ إلى البيانات الجزئية وضمان عدم إفشاء سرية البيانات الشخصية. ويعدّ وجود هذه الترتيبات ووضوحها (سواء أقرها القانون أو الأنظمة الإضافية، أو المراسيم، إلخ) ضرورياً لتعزيز ثقة العامة في الاستخدام السليم للبيانات الجزئية. من البديهي أن الطريقة الفضلى تتمثل في وضع أحكام قانونية. أما في البلدان التي يتعدّر فيها ذلك، ينبغي وضع شكل آخر من أشكال الأحكام الإدارية، كما يتعين على السلطات المؤهلة في مسألة السرية، حيثما وجدت، أن توافق على الأحكام القانونية (أو غيرها) قبل أن يتم إقرارها في القانون. وفي غياب مثل هذه السلطة، تستطيع بعض المنظمات غير الحكومية أن تقوم بمهمة ”الإشراف“ فيما يتعلق بمسائل السرية. وقد يكون من الحكمة جعلها تمثل لأي شكل من أشكال الأحكام القانونية أو غيرها من الأحكام النافذة، أو على الأقل جعلها تستجيب لأي شواغل خطيرة قد تواجهها.

لا تتمتع جميع البلدان بأساس تشريعي. لذا كحد أدنى، ينبغي أن يعطى لكشف البيانات الجزئية دعم أي شكل من أشكال السلطة، ولكن يفضل أن توضع تشريعات تمكّن من ذلك^[24] (ترجمة من الإنجليزية)

المبدأ الرابع: ينبغي ضمان شفافية أنظمة نفاذ الباحثين إلى البيانات الجزئية، وتلك المتعلقة باستخدامات البيانات الجزئية ومستخدامها، وإتاحتها للعامة

”تتمثل أهمية المبدأ الرابع في أنه يطمئن الجمهور بأن البيانات الجزئية تستخدم بحكمة وأن القرارات التي تتعلق بالبيانات الجزئية تتخذ على أساس موضوعي. وفي هذا السياق، يعنى مكتب الإحصاءات الوطني بتحديد ما إذا كان يمكن الإفصاح عن البيانات الجزئية أم لا، تحت أي ظروف، ولأي مستخدم. ومع ذلك ينبغي أن يضمن شفافية هذه القرارات. وفي هذا الصدد، يشكل الموقع الإلكتروني لمكتب الإحصاءات الوطني أداة فعالة لضمان الامتثال للقواعد السارية، وكذلك لتوفير المعلومات عن كيفية النفاذ إلى التقارير البحثية المستندة إلى البيانات الجزئية الموضوعة في التصرف^[24].

عنها، والتي تحتوي على معلومات سرية. (...)

كما أن الامتثال لقواعد السلوك المهنية يمكن أن يسهل احترام المتطلبات القانونية^[17].

المبادئ الأساسية للإحصاء الرسمي في الأمم المتحدة

استندت بلدان كثيرة إلى المبادئ الأساسية للإحصاءات الرسمية للأمم المتحدة في صياغة تشريعاتها، وسيكون من المجدي التطرق إلى هذه المبادئ نظراً إلى علاقتها بسرية الإحصاءات.

ينص المبدأ السادس من المبادئ الأساسية للإحصاءات الرسمية للأمم المتحدة على ما يلي: ”في ما يتعلق بالبيانات الفردية التي تجمعها الجهات المكلفة بذلك لمؤسسة الإحصاءات، سواء أكانت تخص أشخاصاً ماديين أو معنويين، ينبغي أن تكون سرية للغاية ويجب ألا تستخدم لغير أغراض إحصائية.“^[27]

ويجب أن تلبى جميع المبادئ التي تحكم النفاذ إلى البيانات الجزئية هذا الشرط أو أن تخضع لأحكام القانون النافذ في مكتب الإحصاءات الوطني. أما النقاط التالية، فينبغي أن تؤخذ في الاعتبار في إطار إدارة سرية البيانات الجزئية.

المبدأ الأول: الاستخدام السليم للبيانات الجزئية

”يمكن استخدام البيانات الجزئية المجموعة في إطار الإحصاءات الرسمية لأغراض التحليل الإحصائي، من أجل دعم البحوث، بشرط حماية سرية هذه البيانات. (...)

لا يتعارض إيصال البيانات الجزئية إلى الباحثين مع المبدأ الأساسي السادس للأمم المتحدة في حال استحالة تحديد البيانات المتعلقة بالأفراد. ولا يشكل المبدأ الأول المذكور أعلاه شرطاً لنشر البيانات الجزئية. إذ ينبغي أن تعود مهمة اتخاذ القرار بشأن تقديم هذه البيانات الجزئية أو عدم تقديمها إلى مكتب الإحصاءات الوطني. كما يمكن لاعتبارات أخرى (مثل جودة البيانات الجزئية) أن تعيق النفاذ إلى البيانات الجزئية. ويمكن أيضاً أن يكون من غير الملائم تقديم البيانات الجزئية إلى أشخاص أو مؤسسات معينة^[24].

المبدأ الثاني: لا ينبغي إيصال البيانات الجزئية إلا عند انتهاء الإحصاءات

”وفقاً للمبدأ الثاني المذكور، ينبغي التمييز بين استخدام البيانات لأغراض إحصائية أو تحليلية وبين استخدام إداري لهذه الأخيرة. في الحالة الأولى، يتمثل الهدف المنشود في وضع إحصاءات تتعلق بجماعة (أشخاص حقيقيين أو معنويين). أما في الحالة الثانية، فيتمثل الهدف المنشود في إطار استخدام إداري للبيانات في الحصول على معلومات عن شخص مادي أو معنوي من أجل اتخاذ قرار من شأنه أن يعود بالنفع أو الضرر على شخص ما، مثل طلبات الحصول على البيانات الفردية بموجب قرار قضائي. ولغاية كسب ثقة الجمهور في النظام الإحصائي الرسمي، لا تتماشى هذه الطلبات مع هذا المبدأ وينبغي رفضها بشكل منهجي، على الرغم من شرعيتها.

4.3. التعرّض إلى الانتقاد والمعارضة

”نخشى بعض مكاتب الإحصاء الوطنية ألا تكون نوعية بياناتها الجزئية كافية لتنتشر على نطاق أوسع، وإذا تبين أن النوعية جيدة بما فيه الكفاية بما يسمح بتطوير إحصاءات إجمالية، فقد لا تكون كافية أحياناً لإجراء تحليل مفصّل للغاية. في بعض الحالات، تم إجراء تعديلات على إحصائيات مجمعة في مرحلة تحرير النتائج، من دون تغيير البيانات الجزئية. لذا، قد تظهر تناقضات بين نتائج البحوث المستندة إلى البيانات الجزئية وتلك المستندة إلى البيانات المجمعة المنشورة.” [24] أما أجزاء مجموعات البيانات التي لا تعتبر موثوقة بما فيه الكفاية، فيمكن حذفها قبل النشر. كما يتعيّن على منتجي البيانات أن يكونوا منفتحين وشفافين بقدر كاف حيال مسألة الجودة.

من ناحية أخرى، يتمشى إيصال البيانات الجزئية إلى الباحثين مع إمكانية نشر نتائج غير مماثلة لتقديرات الوكالة التي أصدرتها، وإذا كانت الجهة التي أصدرت البيانات وكالة إحصاءات رسمية، يمكن أن يسبب ذلك نزاعاً بين التقديرات المتناقضة (الرسمية وغير الرسمية) ويؤدي إلى التشكيك في البيانات، أو حتى تداعيات سياسية. وقد تساهم عوامل كثيرة في إنشاء مثل هذه الاختلافات. أولاً، قد تكون التقديرات الرسمية مخطئة، وفي هذه الحالة يجدر إجراء مراقبة خارجية، ومن ناحية أخرى، قد تعود الاختلافات إلى استخدام نسخ مختلفة للبيانات (الملف الرئيسي الكامل أو نسخة عامة متناقضة أو محدودة، والتغيرات الأخرى التي أجراها الباحث، إلخ). لذا يفترض بأن تكون الاختلافات طفيفة وأن يتم العثور بسهولة على تفسير لها.

وأخيراً، قد تعود الفوارق إلى استخدام منهجيات مختلفة. وغالباً ما يشكل هذا معضلة أكبر بالنسبة إلى منتجي البيانات، إذ أن الجمهور الكبير لا يحصل يفهم دائماً التفسيرات التقنية المعقدة. لذا من المهم أن يكون منتجو البيانات قادرين على الدفاع عن تقديراتهم. ويستلزم ذلك توثيق عمليات جمع البيانات ومعالجتها وتحليلها من الألف إلى الياء. بالإضافة إلى ذلك، يجب أن يكون النفاذ إلى هذه المعلومات سهلاً. وقد يحدث أن يصيغ النتائج المنشورة خبراء خارجيون، أو بمساعدتهم، من دون أن يبقوا للرد على الأسئلة التي تثيرها هذه النتائج. وفي هذه الحالة، يمكن لمنتجي البيانات أن يستبقوا هذا الاحتمال عن طريق اعتماد وتنفيذ ممارسات توثيق وحفظ دقيقة للغاية تتماشى مع معيار إعادة إنتاج البيانات. ويمكن وصف هذا المعيار في جوهره كما يلي: ” (...) تتمثل الطريقة الوحيدة لفهم تحليل تجريبي وتقييمه في معرفة دقيقة لعملية إنشاء البيانات وتحليلها. (...) ويقضي معيار إعادة الإنتاج بأن تكون المعلومات المتاحة من الجودة بما يكفي لفهم الدراسات السابقة وتقييمها والرجوع إليها في حال كرر طرف ثالث النتائج من دون أي معلومات أخرى من المؤلف.” [10] (ترجمة من الإنجليزية، أنظر أيضاً [11]).

4.4. التكاليف

”يمكن لمكاتب الإحصاء الوطنية أيضاً أن تشغل بمسألة التكاليف، ليس فقط تلك التي تؤدي إلى إنشاء ملفات البيانات الجزئية وتوثيقها، بل أيضاً تلك المتعلقة باعتماد أدوات النفاذ والحماية، والدعم، وترخيص الاستقصاءات التي تجريها جماعة الباحثين، وفي الواقع، يحتاج المستخدمون الجدد لملفات بيانات هذه المكاتب إلى المساعدة للوصول إلى هيكلية الملفات المعقدة وتعريفات المتغيرات. وعلى الرغم من أن مكاتب الإحصاء الوطنية هي التي تتحمل نفقات هذه المسائل، فهي لا تملك عادةً بقدر من الميزانية يكفي لتمويل الأعمال الإضافية التي يجب متابعتها في هذا السياق. أما بالنسبة إلى الباحثين، فهم غالباً لا يملكون

الإطار 8 أمثلة على التشريعات في مسألة السرية

فيما يلي أمثلة عن قوانين الإحصاءات وتدبير بثّ البيانات الجزئية:

1. تدار أنشطة مكتب Census Bureau الأمريكي وفقاً لأحكام القسم 9 من العنوان 31 من قانون الولايات المتحدة. ”موجب أحكام القسم 9 من العنوان 31 من قانون الولايات المتحدة، يمنع نشر أو إيصال أي معلومات من شأنها أن تتيح التعرّف إلى هوية مؤسسة، أو شخص، أو أسرة بالتحديد. وتضم عملية ضبط إفشاء السرية إجراءات تتخذ للتأكد من أن المعطيات المذكورة في العنوان 31 تحضّر قبل النشر، بطريقة تحمي الكشف المتعسف عن الهوية. وتشمل هذه الإجراءات وسائل تحدّد من الكشف عن الهوية وعملية ضبط تضمن أن هذه الطرق المتبعة توفر حماية ملائمة للبيانات.“ (ترجمة من الإنجليزية)

المصدر: <http://www.census.gov/srd/sdc/wendy.drb.faq.pdf>

2. ينص قانون الإحصاءات الكندي على القاعدة التالية: ”لا يجوز لأي شخص معتمد بموجب المادة 6 الكشف أو التسبب بالكشف بأي وسيلة كانت، عن معلومات تم الحصول عليها بموجب هذا القانون بحيث يتيح هذا الكشف أن تلصق بفرد أو شركة أو مؤسسة تفاصيل تم الحصول عليها في مسح يخصهم حصراً.“

المصدر: <http://www.statcan.gc.ca/about-aperçu/act-loi-fra.htm>

تنشر Statistique Canada ملفات بيانات جزئية. أما سياستها لنشر البيانات الجزئية فتتص على أن نشر ملفات البيانات الجزئية يخضع إلى الشروط التالية:

(أ) إذا كان البثّ يحسّن القيمة التحليلية للبيانات المجموعة إلى حد كبير؛
(ب) وشرط أن تتخذ الوكالة كل الإجراءات الممكنة لمنع التعرّف إلى وحدات الملاحظات في الاستقصاء.

3. في تايلاندا، يضم القانون الأحكام التالية: ”يجب التعامل مع المعلومات الشخصية التي يتم جمعها بموجب هذا القانون بمنتهى السرية. ويتعيّن على أي شخص ملتزم بموجب هذه الأحكام أو مسؤول عن حفظ هذه المعلومات الامتناع عن كشفها لأي شخص غير ملتزم بموجب هذه الأحكام، باستثناء الحالات التالية:

(1) إذا كشف عن المعلومات في إطار تحقيق أو ملاحقة قضائية في جريمة.

(2) إذا توجهت المعلومات المكشوف عنها لتستخدم لغايات تحضير أو تحليل إحصاءات أو بحث إحصائي، شرط ألا يلحق هذا الكشف أذى بأصحاب البيانات وأن تبقى هويتهم محفوظة.“ (ترجمة من الإنجليزية)

المصدر: <http://web.nso.go.th/eng/en/about/about0.htm> section 15

الوسائل اللازمة لتمويل جزء كبير من التكاليف“ [24] (ترجمة من الإنجليزية) وبالتالي، ينبغي إدخال هذه التكاليف، بقدر ما أمكن، في ميزانية الاستقصاء، كما ينبغي أن تساهم في الاستفادة القصوى من النتائج التي يتم الحصول عليها. كما أن إيصال النتائج المستخلصة من البيانات يصب في المصلحة العامة، إذ يسمح بتوفير المعلومات لصانعي السياسات والعامة. بالإضافة إلى ذلك، حين تستخدم بيانات الاستقصاء على نطاق أوسع، يحميها ذلك من تقليص حصة اقتطاعات الموازنة للغايات الإحصائية. كما أن الاستقصاءات القليلة الفائدة من ناحية مساعدتها في وضع السياسات العامة مهددة أكثر فأكثر بالتخلي عنها.

4.5 خسارة الحصرية

عند نشر البيانات الجزئية، يفقد مالكوها الحق الحصري في الاطلاع على البيانات. وتمثل هذه الناحية إشكالية بالنسبة إلى الباحثين الجامعيين أكثر مما هي إشكالية بالنسبة إلى منتجي البيانات الرسمية، وإن كان هؤلاء الآخرون (أو بعض الأعضاء في فريق عملهم) يستفيدون أحياناً من حق النفاذ الحصري من أجل عرض خدمات استشارية. وقد أصبح رعاة الاستقصاءات يحددون أكثر فأكثر فترة ”معقولة“ للنفاذ الحصري إلى البيانات للمنتج. وعقب هذه الفتوة، يصبح النفاذ إلى البيانات متاحاً للمستخدمين.

4.6 القدرات التقنية

ثمّة بعض القدرات التقنية اللازمة لتحقيق بث ملفات البيانات الجزئية. يجب أن ترفق الملفات بتوثيق مفصل (ويفضّل أن يكون متماشياً مع معيار DDI للبيانات الجزئية) وأن تكون محفوظة بطريقة ملائمة. كما ينبغي أن تفحص هذه الملفات من أجل تقييم خطر الكشف عن معلومات شخصية، وتقليص هذا الخطر عن طريق اللجوء إلى تقنيات عديدة. وفي هذا الصدد، يفضل الفصل 10 المستلزمات التقنية الخاصة ببث ملفات البيانات الجزئية.

5. إلى من تتوجه هذه البيانات الجزئية؟

تتوجه ملفات البيانات الجزئية إلى المختصين الذين يتمتعون بكفاءات كمية عالية. ومن الأمثلة على ذلك:

- صانعو القرارات السياسية والباحثين الموظفين لدى إدارات التخطيط وأقسامها.
- المنظمات الدولية وغيرها من الرعاة.
- المعاهد الجامعية ومراكز الأبحاث المنخرطة في العمل البحثي الاقتصادي والاجتماعي.
- فريق عمل الجامعات والطلاب.
- غيرهم من المستخدمين المنخرطين في البحث العلمي.

يتمثل أحد أساسات نشر البيانات الجزئية في الإنصاف. فحين يكون النشر ممكناً من الناحية القانونية، يفضل أن يتم اقتراح البيانات الجزئية التي تم الحصول عليها عن طريق عمليات الجمع التي تمّولها الأموال العامة على جميع المستخدمين المحتملين اتباعاً لقاعدة النفاذ المفتوح، مع توفير البيانات الوصفية المكتملة بأكثر قدر ممكن. "ونعني بالنفاذ المفتوح إلى البيانات في ظل شروط مساوية لكل أعضاء الجماعة العلمية الدولية، وبأدنى تكلفة ممكنة، ويفضل ألا تتعدى التكلفة الهامشية للنشر. كما ينبغي أن يكون النفاذ المفتوح إلى بيانات البحث الممول من الأموال العامة سهلاً وسريعاً وسهل الاستخدام، ويفضل أن يكون ذلك عن طريق الإنترنت". [17]

تقترح مكاتب الإحصاء الوطنية عادة منتجات مختلفة، موجهة إلى جمهور مختلف. إذ غالباً ما تتوجه بيانات التلخيص (الجدول، الرسوم الجغرافية، التحليل) إلى الجمهور الواسع. وهي تنشر وتدرج في مواقع المنظمات على شبكة الويب. أما ملفات البيانات الجزئية، فهي موجهة إلى باحثي المؤسسات المختلفة (الوكالات الحكومية والوزارات، المنظمات غير الحكومية، معاهد الأبحاث، الجامعات، المنظمات الدولية) التي تطلق عليها التسمية الشائعة "جماعات الباحثين".

"ماذا نقصد بجماعة الباحثين؟ إنها تشمل بالطبع أولئك الذين يعملون في المؤسسات الجامعية، بل أيضاً الباحثين العاملين في المنظمات غير الحكومية والمنظمات الدولية. ومن ناحية أخرى، تشمل بعض الباحثين الذين يودون النفاذ إلى البيانات الجزئية في قلب وكالات أو مؤسسات تمّولها الحكومة". [24]

تنص قوانين إحصاءات عديدة على وجوب إتاحة البيانات لأغراض إحصائية مشروعة ولدعم البحث. وغالباً ما يعدّ الاستخدام التجاري في إطار تعاقد غير ملائم مع هذه القاعدة. إذ يشكلّ التشديد على الاستخدام المحتمل للبيانات الجزئية قاعدة أساسية للائتمان إلى القوانين التي تخضع لها أنشطة المنظمات غير الحكومية.

ويصف المكتب الإحصائي للاتحاد الأوروبي، أوروبستات Eurostat، النفاذ إلى البيانات الجزئية كما يلي:

"تتاح البيانات الجزئية في المقام الأول إلى جماعة العلماء أو المعاهد أو الجامعات، لأغراض البحث حصرياً وبعد تقديم مشروع بحث. وفي حالات نادرة للغاية، يوسّع حق النفاذ ليشمل طلاباً أو جمهوراً غير علمي أوسع". أما مجموعة البيانات الجزئية المغفلة التي توضع في تصرف الجمهور الكبير فتسمى الملفات المتاحة للاستخدام العام. وإذا كانت تنشرها المعاهد الإحصائية على نطاق واسع

مجاناً، فهي تلقى في المقابل اهتماماً محدوداً من حيث ثبات المعلومات بالنسبة إلى الأنشطة العامة. "وبشكل عام، لا يحق للمؤسسات الخاصة العمل ببيانات جزئية". [6]

وإذا كان بثّ البيانات الجزئية يتوجه عادة بشكل أساسي إلى الباحثين، فهم يتمتعون ببعض المزايا في معظم الأحيان. إذ يتم تعريفهم على أنهم مستخدمون أهل للثقة، أي باحثون حقيقيون يعملون بنية سليمة.

وفي الحقيقة، لا تنقل "ملكية" ملفات البيانات إلى الباحثين: بل يستفيد هؤلاء من رخصة استخدام غير قابلة للنقل. ويسمح ذلك لمكتب الإحصاءات الوطني بتحديد الهدف المنشود. كما يفضل أن يضع مكتب الإحصاءات الوطني سياسة تسمح له بصياغة شروط النفاذ إلى هذه البيانات.

يجدر أن تكون هذه السياسة عامة بأقصى قدر ممكن وأن تضم معلومات لكل أنواع المستخدمين. ويمكن أن تفرض شروط معينة على بعض المستخدمين بحسب جنسيتهم والمنظمة التي ينتمون إليها. ومن الأمثلة على ذلك:

المستخدمون الوطنيون:

- يمكن للممثلين الرسميين المعيّنين صراحة بالتشريعات التي تحكم الإحصاءات أن يستفيدوا من حق نفاذ مسهل إلى البيانات الجزئية الرسمية، وهم يخضعون إلى القواعد نفسها التي يخضع إليها مكتب الإحصاءات الوطني. كما يمكن أن يكونوا معتمدين وأن يكون تطبيق العقوبات أسهل، ما يصب في أهداف الحكومة.

- يحق للمستخدمين الوطنيين العاملين بنظام تشريعات الإحصاءات على الرغم من ذلك استخدام ملفات البيانات، شرط أن يوقعوا عقد ترخيص ملائم.

المستخدمون الدوليون:

- في حالات كثيرة، ثمة التزام بإيصال البيانات إلى منظمات دولية. وقد يرتبط هذا الالتزام بانتماء الوكالة إلى مجموعة دولية أو بعقود تمويل لمشاريع مهمة، أو حتى بمشاركة بلد معين بمشاريع التنمية الدولية.

- يشكل منح تراخيص استخدام إلى باحثين يعملون في جامعات أو مراكز أبحاث أجنبية مسألة أكثر حساسية. وفي الواقع، يصعب في هذه الحالة الإلزام باحترام الأحكام المطبقة في هذه البلدان. ولكن قد يكون من المجدي مشاركة البيانات مع هؤلاء الباحثين، فهم يتمتعون بخبرة مجمعة غنية. ويمكن تقليص الخطر عن طريق منح الترخيص إلى مؤسسة (جامعة مثلاً) عوضاً عن منحها إلى فرد بالتحديد (اقرأ المعلومات المكتملة المقدّمة لاحقاً).

المساعدون التقنيون:

يستفيد فريق عمل مكتب الإحصاءات الوطني من امتحان للمستشارين، أجنب كانوا أم لا، لمعالجة و/أو تحليل بيانات الاستقصاء. حين يكون هذا الإجراء متماشياً مع أهداف مكتب الإحصاءات الوطني، ينصح بذكره في سياسة النشر، إلا

6 يمثل الموقع الرسمي لـ Eurostat منصة للنفاذ إلى البيانات الجزئية.

ويسمح الرابط أدناه بالاطلاع على إجراءات السرية المطبقة في Eurostat:

http://epp.eurostat.ec.europa.eu/portal/page/portal/research_methodology/statistical_confidentiality/confidential_data/introduction

، تمت مراجعة الموقع في 10 يناير/ كانون الثاني 2011

iv. عند ارتكاب المخالفات، رفض تقديم أي خدمة إلى الباحث ومن خلاله إلى المؤسسة التي ينتمي إليها لفترة معينة (حتى تتخذ المؤسسة الإجراءات التأديبية بحق الباحث المخالف مثلاً). ومن الضروري للغاية أن يفهموا أن جزءاً من الجمهور قد يراهن على كشف البيانات الجزئية في المستقبل إلى جماعة الباحثين. لذا من الملائم في هذا الحال إحالة الأمر إلى القضاء.” [24] (ترجمة من الإنجليزية)

حين يقدم باحث طلباً للنفاذ إلى البيانات نيابة عن مؤسسة، يفضل أن يقوم بذلك باسم هذه المؤسسة (أو الوكالة، أو الشركة التي يعمل لديها، إلخ) وليس بصفته فرداً. إذ تتمتع هذه المؤسسات بسمعة تودّ الحفاظ عليها، ما يحث على احترام الالتزامات المتعهد بها. غير أنه من الصعب تطبيق هذه العقود في حالة طلب النفاذ إلى البيانات عبر الحدود. وقد تتمثل أحد الحلول لهذه المشكلة في تعاون أرشيفات البيانات أو مكاتب الإحصاء الوطنية في البلد الذي قدم الطلب.

إذا وقع المستشارون العقد نفسه الذي يوقعه الباحثون الآخرون. وتفضل بنود هذا العقد ألا يتم إيصال البيانات من دون موافقة المنتج المسبقة. كما يجدر فرض تعهد بالسرية.

يمكن لمكاتب الإحصاء الوطنية أن تنقل جزءاً من المخاطر إلى الباحثين عن طريق اتخاذ الإجراءات التالية:

i. ”أن يطلب منهم البرهنة عن شرعيتهم بصفتهم باحثين، وإظهار فائدة أبحاثهم بالنسبة إلى العامة وتقديم الأدلة التي تبين أن استخدام البيانات الجزئية مهم في هذا الإطار؛

ii. أن يوقع الباحثون تعهداً قانونياً ملزماً ينص على عقوبات مثيلة بتلك التي تطبق على موظفي مكتب الإحصاءات الوطني إذا انتهكوا القواعد التي تتعلق بالسرية؛

iii. تفسير أسباب الحذر الذي تبديه مكاتب الإحصاء الوطنية. التأكد من أن الباحثين يعرفون تمام المعرفة التزاماتهم عن طريق إبلاغهم بها بالطريقة السليمة. وضع تدابير المتابعة والإشراف الفعالة. وقد يكون من المجدي وضع قواعد للسلوك المهني بالتعاون مع جماعة الباحثين؛

الإطار 9 مثال عن التعهد بالسرية

ألتزم بـ:

1. ألا أقوم بأي نسخة جزئية أو كاملة يمكن أن يرخص لي فريق العمل في مركز البيانات الوطني المحمي (المشار إليه فيما يلي باسم ”المركز“) بالنفاذ إليها. وأنا على علم بعدم جواز إخراج أي من البيانات أو المعلومات الشخصية التي تتم مراجعتها أو يتم الحصول عليها بموجب ذلك من المركز إلى خارجه.
2. أن أعيد إلى فريق العمل في المركز مجموع الوثائق السرية التي سيضعها المركز في عهدي خلال أبحاثي، وكل وثيقة مطلوبة.
3. أن لا أقوم بأي شيء يسمح لي بالتعرف على هوية شخص أو مؤسسة، أو وحدة في العينة لم تذكر هويتها في ملفات البيانات المتاحة للاستخدام العام.
4. أن أحافظ على أكبر قدر من السرية حيال هوية المؤسسات أو الأشخاص التي تكتشف عن طريق الخطأ، سواء أكان ذلك بواسطة محادثة أو في إطار التحليل. ويتم إبلاغ فريق العمل في المركز على الفور بأي اكتشاف غير إرادي للهوية في إطار أعمال التحليل التي أجريها.
5. أن لا أدون انطباعات، أو أنشئ ملفات إلكترونية أو وثائق أو أي مستند دون أن يخضعها فريق العمل في المركز لفحص دقيق مسبق، من أجل استبعاد أي خطر الكشف عنها. وأنا على علم بأن المركز سيجري ضبطاً للكشف، وأن أي سحب للبيانات، سواء أكان بصيغة إلكترونية أو بنسخة ورقية، سيخضع لترخيص مسبق من قبل المركز.
6. أن لا أنقل أي معلومة مدوّنة باليد عن هوية مؤسسة أو شخص أو عن معلومات الاتصال الجغرافية التي يتم التوصل إليها في إطار أبحاثي في المركز.
7. أن أعتد سلوكاً يحترم مبادئ قواعد السلوك الحسن التي تعدّ ملائمة لمؤسسة بحث علمي.

لقد أبلغت بأن كل مخالفة متعمدة لأحد هذه الشروط، سترتب عليها إلغاء ترخيص النفاذ إلى البيانات. كما أنني على علم بأن هذه المخالفة قد تغلق أمامي كل أبواب المركز نهائياً للفترة التي يقدر المدير بأنها مناسبة لحماية نزاهة المركز وسريته.

التاريخ

اسم المستشار وتوقيعه

التاريخ

اسم الشاهد وتوقيعه

يستوحي هذا التعهد بالسرية أحكامه من نسخة عام 2008 للعقد الذي يحكم شروط النفاذ إلى البيانات في مركز أبحاث البيانات Research Data Center، ومركز إحصاءات الصحة الأمريكي (NCHS) Conditions of Access to Confidential Data in the Research Data Center of the US National Center for Health Statistics (Agreement Regarding (16)

6. ما هي الشروط التي تحكم بثّ البيانات؟

لجنة ضبط الكشف عن البيانات الجزئية، المكلفة بالمهام التالية:

- فحص كل طلبات نشر ملفات البيانات الجزئية المغفلة التي يقدّمها مدراء الاستقصاءات، وفقاً لمعايير موضوعة مسبقاً؛
- الموافقة على نشر كل الملفات أو الإشارة إلى مدراء الاستقصاءات بالطرق التي تسمح لهم بتحسينها؛
- الإشراف على منح التراخيص وتقليص المخاطر المرتبطة بمخالفات محتملة؛
- مراجعة الخطوط الإرشادية التي يطبقها مدراء الاستقصاءات، في ما يتعلق بملفات البيانات الجزئية المغفلة.

لجنة البثّ، المكلفة بالمهام التالية:

- دراسة طلبات النفاذ التي يقدمها الباحثون المتعلقة بالبيانات الجزئية الخاضعة لترخيص؛
- إتاحة النفاذ إلى ملفات البيانات إلى المستخدمين المرخصين؛
- والاستجابة إلى طلبات المساعدة والمعلومات المكتملة التي يقدمها المستخدمون.

يوافق المدير العام للمؤسسة على كل إيصال لملفات البيانات الجزئية المغفلة إلى المستخدمين، استناداً إلى نصائح لجنة الكشف وتوصياتها.

- وصف عام لسياسة تكاليف المؤسسة. وفي الأحوال المثالية، تولى هذه السياسة المقام الأول لأوسع نطاق ممكن لاستخدام البيانات التي تنتجها عن طريق جعل النفاذ إليها ممكناً، لذا يتعيّن على منتجي البيانات أن يحرصوا على إدخال تكاليف إنشاء بيانات الملفات الجزئية المغفلة في ميزانية الاستقصاءات. وعلى نحو مواز، يحق لمكتب الإحصاءات الوطني أن يسعى بشكل قانوني إلى استرداد التكاليف المتعلقة بتقديم خدمات خاصة متوجهة حصراً إلى فئة من المستخدمين دون غيرهم.

كما نجد معلومات أكثر تفصيلاً عن التدابير والشروط الخاصة بنشر ملفات البيانات الجزئية في البروتوكولات وقواعد التدابير، لاسيما:

- قواعد طلب النفاذ إلى البيانات (الطلبات على شبكة الإنترنت، نماذج الطلبات المعتمدة، إلخ)
- الترخيص والمنع لمجموع البيانات من مختلف الأنواع (أنظر القسم 6.2)
- الهيئة المسؤولة عن منح تراخيص النفاذ وغيرها من المعلومات العملية بشأن تدابير الضبط
- طرق ضبط الكشف الإحصائي المعتمدة
- المعلومات التي ينبغي أن يقدّمها الباحثون والوجهة المحتملة لهذه المعلومات
- سياسة التكاليف المفصلة
- شكل المساعدة التقنية التي يستفيد منها المستخدمون ونطاقها
- معلومات أخرى عملية

في هذا الإطار، نشرت UK Statistics Authority دليلاً للممارسات السليمة في مجال الإحصاءات الرسمية [22]. وعلى الرغم من أن هذا الدليل

تتمثل مهمة معظم منتجي البيانات التي تمّولها الأموال العامة في بثّ البيانات التي يجمعونها على أوسع نطاق ممكن. غير أن كل هذه المؤسسات تخضع إلى التزامات أخلاقية وقانونية تفرض عليها تدابير تقيّد النشر. "من الضروري للغاية نشر الإحصاءات المفيدة إلى أكبر عدد من الأشخاص، غير أن الأهمية نفسها تولى إلى أن يتم ذلك وفقاً لتدابير لا تضرب بالمؤسسات التي أنتجتها". [15]

"(لذا) ينبغي وضع مجموعة من الإجراءات القانونية، والإدارية والتقنية لكسب ثقة الجمهور والمحافظة عليها". [24]

يجب أن تكون الشروط التي تحكم نشر البيانات الجزئية رسمية وشفافة. وفي الأحوال المثالية، تحدد في سياسة وفي قواعد إجرائية وبروتوكولات مختلفة. وفي هذا السياق، يضم هذا الفصل معلومات عن صياغة هذه السياسة والإجراءات. ولكن ينبغي التنويه بأن هذه المعلومات لا تشكل معياراً يوصى به. بل يعود لكل منتج بيانات أن يحدد سياسته الخاصة وإجراءاته الخاصة، مع مراعاة النواحي التقنية والقانونية والأخلاقية في آن واحد.

عادة ما تحتوي سياسة النشر على أحكام ذات طابع عام، لا سيما:

- موضوع السياسة التي تهدف إلى "تحديد نوع بيانات الملفات الجزئية المعدة للنشر، والأهداف التي يمكن أن تستخدم لأجلها، والشروط التي تحكم وضعها في التصرف".
- بيان موجز بالمبادئ، تفسّر من خلاله المؤسسة لماذا تعدّ بثّ البيانات الجزئية مهماً (قراءة الفصل الثالث).
- التشريعات النافذة في ما يتعلق بمسألة بثّ البيانات الجزئية (أنظر القسم 6.1).
- المبادئ العريضة التي تحكم الحفاظ على خصوصية البيانات وسريتها.
- الأجل المحدد لوضع البيانات في التصرف. مثال: "إن مكتب الإحصاءات الوطنية، إذ يأخذ بالاعتبار أهمية حاجات المستخدمين وأهمية وضع البيانات الحديثة في تصرفهم، سيجهد في سبيل نشر ملفات البيانات الجزئية في مهلة تتراوح بين 6 و12 شهراً، اعتباراً من أول نشر لبيانات الإحصاء المعني. كما يتم أحياناً نشر بيانات الاستقصاءات على مراحل عدة، بحيث يتمكن مكتب الإحصاءات الوطني من ضبطها وتحليلها قبل أن يتم إغفالها لتستخدمها المؤسسة لاحقاً". (أنظر <http://www.surveynetwork.org/home/index.php?q=tools/dissemination>)
- الوظائف والمسؤوليات الأساسية في ما يتعلق بتحديد ووضع سياسة النشر وأنظمتها. ومن الأمثلة على ذلك:

مدراء الاستقصاءات، المكلفون بالمهام التالية:

- تحديد الحاجات الأساسية للمهتمين والحرص على إنشاء ملف مغفل يسمح بالاستجابة إلى حاجات جماعة المستخدمين مع احترام تشرّيعات الإحصاءات؛
- وبدء مراقبة أولى ملف البيانات الجزئية، وتحديد المشاكل المحتملة التي يتوجب حلها، وتقديم نسخة مبدئية إلى لجنة الكشف عن البيانات الجزئية.

تحميل البيانات المرجوة.

يقدم موقع الاحصاء الكندي مثالاً عن ملف متاح للاستخدام العام. وهو ملف بيانات جزئية ناتج عن الاستقصاء المشترك بين كندا والولايات المتحدة بشأن الصحة *Enquête conjointe Canada/Etats Unissur la santé*⁷ (أنظر أيضاً *European Social Survey*⁸). للحصول على ملفات البيانات الجزئية، ليس على المستخدمين إلا التسجيل. فالهدف يتمثل في أن يتم إبلاغهم بالتعديلات المستقبلية التي تجرى على الملفات الموجودة وإتاحة ملفات جديدة.

يستلزم نشر ملفات البيانات الجزئية احترام بعض القواعد المبادئ. وفي هذا الصدد، يحدد الإطار 10 الشروط الأساسية التي تطبق على ملفات البيانات الجزئية المتاحة للاستخدام العام.

6.3 الشروط التي تطبق على الملفات الخاضعة لترخيص

يجب أن تستعيد الشروط العامة التي تتعلق بملفات البيانات الجزئية الخاضعة لترخيص المبادئ العريضة الأساسية، وأن تستكملها أحكام تطبيق على المؤسسة التي ينتمي إليها الباحثون. وثمة احتمالين في هذا الصدد: أولاً، أن تقدم البيانات إلى باحث أو فريق باحثين لغايات محددة؛ وثانياً، أن توضع بتصرف مؤسسة لاستخدام خارجي استناداً إلى عقد يشكل إطاراً لذلك (منظمة دولية أو مركز أبحاث على سبيل المثال). وفي الحالتين، يجب ذكر المؤسسة التي ينتمي إليها الباحثون ويتعين على الممثلين الملائمين أن يوقعوا رخصة الاستخدام.

الترخيص الممنوح إلى باحث أو فريق باحثين لغايات محددة

حين تبث البيانات من أجل مشروع بحث محدد، يجب أن يتم التعريف بفريق البحث. ويكون ذلك عن طريق طلب رسمي للنفذ إلى البيانات (أنظر النموذج المبيّن في الملحق 1). وتنص شروط الحصول على البيانات (أنظر المثال في الإطار 12) على أن الملفات يجب ألا تنقل لأي شخص من خارج المؤسسة وأنها يجب أن تحفظ بطريقة آمنة. كما ينبغي أن تكون غايات الاستخدام مطابقة بقدر الإمكان للقائمة النتائج المرتقبة وسياسة النشر الخاصة بالمؤسسة. كما أن الحق في النفاذ إلى مجموع البيانات الخاضعة لترخيص ينحصر في المؤسسات المعتمدة قانونياً مثل الرعاة (وزارة التخطيط، الجامعات، مراكز الأبحاث، المنظمات الوطنية أو الدولية).

العقد الإطاري المبرم مع مؤسسة

في هذه الحال، من المتفق عليه أن تحفظ البيانات للاستخدام الداخلي في المؤسسة، وللغايات التي تجدها مفيدة، وأن تتخذ الإجراءات الأمنية المناسبة. كما ينبغي ضمان الامتثال للمبادئ النافذة، وأن يحدد شخص معين باسمه ليتحمل رسمياً المسؤولية. وينبغي إبلاغ كل مستخدم بالشروط العامة المطبقة على ملفات البيانات؛ وقد يطلب في هذا السياق تعهد بالسرية. وفي وجود مثل هذا العقد

ليس متخصصاً في نشر ملفات البيانات الجزئية، إلا أنه يشكل مثالاً جيداً على المبادئ والبروتوكولات المصاغة بشكل واضح.

6.1 الأساس التشريعي

إن الأساس التشريعي ضروري لأسباب عدة:

i. لكسب ثقة الجمهور في الأحكام الموضوعية، وبالأحرى ليعلم الجمهور أن ثمة قيوداً قانونية تحدد ما هو مسموح وما هو ممنوع؛

ii. لإنشاء تفاهم متبادل بين مكاتب الإحصاءات الوطنية والباحثين في ما يتعلق بهذه الأحكام؛

iii. لتحقيق قدر أكبر من التعاون في مجال معالجة مشاريع البحث؛

iv. ولوضع أسس نظام يهدف إلى معاقبة المخالفات". (24)

ثمة بعض مكاتب الإحصاءات الوطنية التي لا تستطيع الرجوع إلى أي أحكام قانونية بشأن الإحصاءات لنشر ملفات البيانات الجزئية، كما أن هذه العملية ممنوعة صراحة في بعض الأحيان. وفي قوانين أخرى، لا يتم تناول الموضوع صراحة، بل يترك التقدير إلى المؤسسات المعنية. ونواجه هذه الحالة بشكل خاص حين يعود وضع القوانين إلى زمن لم يكن فيه هذا الاحتمال وارداً. لذا يمكن أن يخضع نشر ملفات البيانات الجزئية إلى مراجعة مسبقة للتشريعات النافذة. فالقانون الكندي الذي يتناول الإحصاءات على سبيل المثال يعود لبداية القرن العشرين، ولكن تمت مراجعته عام 1971، هو وغيره من القوانين لتأهيل مكاتب الإحصاء الوطنية لنشر ملفات البيانات الجزئية.

ومع ذلك، "يجب أن تذكر هذه الأحكام في نص قانوني. ويمكن أن تحدد التفاصيل.../... بسهولة عن طريق الأنظمة، والمراسيم، إلخ، التي لا يقل تأثيرها القانوني عن تأثير القانون. وبغياب الأحكام القانونية، من الضروري إعطاء شكل من أشكال الترخيص. إذ يخاطر مكتب الإحصاءات الوطني بسماعته إذا لم يكن ثمة أي نوع من أنواع السلطة التي تمارس على كشف البيانات، حتى المغفلة. ويجدر بالتشريعات (أو الترخيصات) أن تراعي النواحي التالية:

i. ما المسموح والممنوع، ولأي غايات؛

ii. شروط الكشف؛

iii. والتبعات المترتبة على عدم احترام هذه الشروط" (24)

6.2 الشروط المطبقة على ملفات البيانات الجزئية المتاحة للاستخدام العام

في المبدأ، يمكن لأي شخص قادر على دخول الموقع الإلكتروني الخاص بمركز إحصاءات وطني أن يطلع على البيانات العامة. ولكن يوصى بتحديد شروط لاستخدام هذه البيانات، وإجراءات وقائية ينبغي اتخاذها. وقد لا تتمتع هذه الأحكام بقوة قانونية دائماً، غير أنها تسمح مع ذلك بتوعية المستخدم تجاه المسائل المتناولة. ويمكن أن تنص "شروط الاستخدام" على وجه الخصوص على أنه يمنع القيام بأي محاولة لإعادة اقتطاع المعلومات مع مصادر أخرى. ويتعين على المستخدم أن يقبل على شبكة الانترنت هذه الشروط حتى يتمكن من

7 Statistique Canada، أنظر أيضاً <http://www.statcan.gc.ca/pub/8230022x/2003001/4069119-fra.htm> ، تمت زيارة الموقع في 01 يناير/ كانون الثاني 2011.

8 European Social Survey، أنظر أيضاً http://www.europeansocialsurvey.org/index.php?option=com_content&task=view&id=78&Itemid=190 ، تمت زيارة الموقع في 22 أغسطس/ آب 2007 (متوفر فقط بالإنجليزية).

ونظام أمن ملائم، لا يطلب إتلاف البيانات بعد استخدامها. ويحتوي الإطار 13 على مثال على صياغة عقد إيطاري.

6.4. شروط خاصة بمراكز البيانات المحمية

أنشئت مراكز البيانات المحمية للبيانات الفائقة الحساسية أو المفصلة للغاية التي لا يمكن أن تكون مغفلة بما يكفي للسماح بنشرها نشرًا عامًا.

كما أننا نصادف مصطلحي "مختبر البيانات" أو "مركز بيانات البحث". ويمكن أن ينتقل مركز بيانات محمي إلى مقر مكتب الإحصاء الوطني أو إلى مواقع كبرى مثل الجامعات القريبة من جماعة الباحثين. إذ يسمح ذلك بإتاحة ملفات بيانات كاملة إلى الباحثين من دون التعرض إلى خطر الكشف عن معلومات سرية. وبشكل عام، يشرف فريق عمل خاص بمكتب الإحصاءات الوطني على النفاذ إلى البيانات واستخدامها، كما أن أجهزة الحاسوب لا ينبغي أن تتصل بشبكة خارجية، ناهيك عن أن نتائج أعمال الباحثين تخضع بالضرورة إلى تحليل مكتب الإحصاء الوطني. وتهدف هذه الخطوات كلها إلى ضمان احترام سرية البيانات. وفي هذا السياق، يقدم الملحق 2 نموذجاً عن سياسة نفاذ معتمدة في مركز بيانات محمية. أما الملحق 3 فيعرض بالمقابل نموذجاً نمطياً لطلب نفاذ إلى مركز بيانات محمي.

وتتميز مراكز البيانات المحمية بأنها تقدّم بيانات جزئية مفصلة، غير أن السمة السلبية تتمثل في أنها تفرض على الباحثين مكان عمل مختلف عن مكان عملهم.

كما أن تكلفة إنشاء مركز كهذا واستغلاله مرتفعة. ومع ذلك، فقد اختارت بلدان عديدة أن تتيح للباحثين إمكانية النفاذ إلى البيانات الجزئية ميدانياً، أما هؤلاء الباحثون فمعتمدون بما يتماشى مع القوانين الإحصائية، شأنهم شأن الموظفين الدائمين في مكتب الإحصاءات الوطني. وغالباً ما يشجع هذا النظام الباحثين الذين يعيشون على مقربة من مكتب الإحصاءات الوطني.

6.5. إدارة مخالفات الباحثين

تظهر تجارب البلدان في إطار ممارسات بثّ البيانات الجزئية أن حالات انتهاك سرية ملفات البيانات نادرة للغاية، إذ تسيئ هذه المخالفات إلى مصالح الباحثين. فهي تضع على المحك سمعتهم الشخصية وسمعة المؤسسة التي ينتمون إليها، ولكن مع ذلك، ينبغي الاعتراف بوجود هذه الانتهاكات، سواء أكانت متعمدة أو عرضية.

نظراً إلى أهمية اعتماد سياسة نشر صالحة للبيانات الجزئية، تضع مكاتب الإحصاءات الوطنية مسبقاً تدابير لتنفيذ أحكام مناسبة:

- يتعيّن إيقاف أي انتهاك على الفور، وهذا أمر ضروري للحفاظ على ثقة المستجيبين ومصادقية المؤسسة؛
- في حال مخالفة القانون، يجب النظر في ملاحقة الموضوع قضائياً؛
- إذا لم يحترم الباحثون التزاماتهم، يمكن أن يلجأ مكتب الإحصاءات الوطني إلى تعليق حقوق النفاذ للأفراد والمؤسسة؛
- إذا قامت مؤسسة ممثلة باحث بالالتزام التعاقدية، تتكفل هي بالنظر في العقوبات التي تتخذها بحق أحد أفراد فريق عملها عوضاً عن مكتب الإحصاءات الوطنية، ويمكن أن يؤدي فقدان الحق في النفاذ الفردي إلى تطبيق العقوبة نفسها على كامل المؤسسة؛
- يتخذ مركز الإحصاءات الوطنية الإجراءات اللازمة لمنع انتهاكات أخرى؛
- أما بالنسبة إلى الانتهاكات الطفيفة، فيكفي توجيه إنذار.

الإطار 10 شروط النفاذ والاستخدام المطبقة على ملفات البيانات الجزئية المتاحة للاستخدام العام

1. لا يمكن بأي حال من الأحوال بيع البيانات والوثائق الأخرى التي يقدمها مكتب الإحصاءات الوطني من جديد أو إعطاؤها إلى أشخاص، أو مؤسسات، أو منظمات أخرى من دون موافقة خطية مسبقة من طرف مكتب الإحصاءات الوطني.
2. تستخدم البيانات حصراً لغايات إحصائية أو في إطار أعمال بحث علمي. كما أنها تهدف فقط إلى تشكيل بيانات مجمعة، بما في ذلك تصميم المعلومات. أما البحث عن أشخاص أو منظمات بحد ذاتها فمحظور بجميع أشكاله.
3. تحظر أي محاولة لإعادة تحديد المستجيبين. وعلاوة على ذلك، يمنع استخدام هوية أي شخص أو منظمة تم اكتشافها مصادفةً، وينبغي إبلاغ مركز الإحصاءات الوطني بأي اكتشاف من هذا النوع.
4. تحظر أي محاولة لإعادة اقتطاع مجموع البيانات التي أتاحتها مكتب الإحصاءات الوطني، سواء مع بيانات مكتب الإحصاءات الوطني أو غيرها من مجموعات البيانات التي من شأنها أن تسمح بالتعرّف إلى هوية أشخاص أو منظمات.
5. يجب أن تحتوي كل الكتب أو المقالات أو وثائق العمل أو الدراسات أو الأطروحات أو التقارير أو غيرها من المنشورات المستندة إلى البيانات التي قدّمها مكتب الإحصاءات الوطني على مراجع لمصدرها، وفقاً لطلب ذكر مراجع الاقتباسات الخاص بمجموع البيانات المقدّمة.
6. تقدّم إلى مكتب الإحصاءات الوطني نسخة بصيغة إلكترونية تضم كل المنشورات المستندة إلى البيانات المطلوبة.
7. لا تتحمل الجهة التي جمعت البيانات في المرحلة الأولى، ولا مكتب الإحصاءات الوطني، ولا الجهات الممولة المعنية بأي شكل من الأشكال المسؤولية المرتبطة باستخدام هذه البيانات، أو تحليلها، أو النتائج المستخلصة منها.

ملاحظات:

- تشير النقطتان 3 و6 إلى أن المستخدمين يتمتعون بإمكانية التواصل بسهولة مع منتج البيانات. ومن المجدي ذكر رقم الهاتف، وعنوان البريد الإلكتروني، وإذا أمكن، فتح نظام لإبداء الآراء على شبكة الإنترنت.
- للاطلاع على النقطة 5، أنظر الإطار 11.

الإطار 11 كيف يتم ذكر ملف بيانات إلكتروني

لحل هذه المشكلة (وغيرها من المشاكل)، قدّم ألتمان وكنينغ & G. King & M. Altman (من مركز بيانات تابع لجامعة هارفرد ومعهد ماسشوستس التقني Harvard-MIT) الاقتراح التالي: "يجب أن تشمل المراجع التي تشير إلى البيانات الإلكترونية ستة عناصر أساسية. تعدّ العناصر الثلاثة الأولى تقليدية؛ إذ نجدها هي نفسها في الوثائق الورقية. وهي تتمثل في الكاتب (أو الكتاب)، وتاريخ النشر، وعنوان مجموعة البيانات. وينبغي أن يتكّيف تقديم هذه العناصر مع المقالات أو الدراسات التي تذكر فيها المرجعية. إذ يسمح اسم الكاتب، والتاريخ، والعنوان بمعرفة نوع البيانات المعنية ومعرفة متى تمت مراجعتها. غير أن هذه العناصر لا تكفي لتحديد مجموعة بيانات بطريقة أحادية، ولا لتحديد موقعها، ولا للحصول عليها أن التحقق منها بطريقة موثوقة. وبالتالي، اخترنا أن نضيف ثلاثة عناصر أخرى عن طريق استخدام تقنيات عصرية. وقد صُمم كل عنصر بحيث لا يفقد صلاحيته في حال حدوث تطور تكنولوجي. كما يتمثل الهدف أيضاً في الاستفادة القصوى من الصيغة الإلكترونية للبيانات الكمية.

يتمثل العنصر الرابع في معرّف لا نظير له. وهو يؤدي دور تعيين اختصاري أو سلسلة من السمات الفريدة والخاصة التي تسمح بالتعرّف إلى مجموعة البيانات على المدى الطويل بشكل مستقل عن مكان تخزينها. (...) إذ تضمن المعرّفات الأحادية ثبات الرابط بين المرجعية والبيانات المعنية. ولكن يتعيّن أيضاً ضمان عدم تغيّر البيانات المعنية بطريقة كبيرة في حال تغيير صيغة الحفظ والتأكد من ذلك. لذا نضيف بصمة إلكترونية عالمية (universal numeric fingerprint، UNF) أما التوقيع الإلكتروني العالمي UNF فهو سلسلة من سمات أبجدية ورقمية ذات طول ثابت، وهي تختصر محتوى مجموعة البيانات. وإن أي تعديل في البيانات، مهما كان طفيفاً، يستوجب توقيعاً إلكترونياً عالمياً جديداً. (...) أما العنصر الأخير في المرجعية المعيارية فهو مدير الوسائط (...) وفي ما يلي مثال عن مرجح كامل يشمل هذا العناصر الأساسية المعيارية:

Social and Political Participation Data. hdl:1902.4/00754 ;
Sidney Verba. 1998. U.S. and Russian
UNF :4:ZNQRI14053UZq389x0Bffg?==;
(ترجمة من الإنجليزية) [1]. "http://hdl%3A1902.4%2F00754".
http://id

ليس من معيار عالمي لذكر مجموع البيانات الجزئية. ولكن يجب أن تحتوي صياغة المراجع على العناصر الأساسية التالية: اسم منتج البيانات، عنوان مجموع البيانات وسنة المرجعية (يليه إشارة تحدد أنها بيانات جزئية)، رقم المرجعية (وهو يحتوي في الحالات المثالية على رقم النسخة) // اسم موزع البيانات وتاريخ الحصول على الملفات (أنظر أيضاً (18) لمزيد من المعلومات حول ذكر المنتجات الإحصائية).

المثال 1

SURVEY (ACS) U.S. Dept. of Commerce, Bureau of the Census.
AMERICAN COMMUNITY: عينة البيانات الجزئية المتاحة للاستخدام
العام (PUMS)، 2005 (ملف معلوماتي). ICPSR04587-v1. واشنطن
العاصمة: U.S. Dept. of Commerce, Bureau of the Census (المنتج)،
Ann Arbor، 2005، معهد ماسشوستس: اتحاد أبحاث العلوم الساسية
والاجتماعية بين الجامعات (ICPSR) (الموزع)، 08-08-2007.

المثال 2

ملف بيانات جزئية متاح للاستخدام العام: استقصاء عن حركية العمل
والدخل، الموجة 8، 2000. الأسر الاقتصادية (ملف معلوماتي) // كندا،
Statistique Canada، قسم استقصاءات الأسر، النسخة 2، أوتاوا، أونتاريو.
مبادرة إتاحة البيانات (الموزع)، 28-08-2003.

غير أن هذه المراجع، على الرغم من صلاحيتها، إلا أنها لا تستجيب لكل المتطلبات ولا تعدّ كافية بالنسبة إلى جميع مراكز البيانات الجامعية. "يمكن تقديم روابط URL، غير أن مدة صلاحيتها عادة ما تكون محدودة. (...) كما أن نسخ البيانات المراجعة غالباً ما تنشر تحت الاسم نفسه، من جون تحديد معياري لرقم النسخة. كما أن المراجعين لا يتبعون قواعد ثابتة، أو بالأحرى لا يتبعون أي قاعدة. وفي هذه الحالة، تذكر المصادر في القسم المخصص للمراجع، أو تدمج في النص أو لا تذكر على الإطلاق. وأياً كان الحال، فإن هذه المراجع نادراً ما تحتوي على القدر الكافي في المعلومات لضمان النفاذ إلى مجموع البيانات نفسها في المستقبل". [1] (ترجمة من الإنجليزية)

معلومات مكتملة عن إدارة انتهاكات عقود الاستخدام:

”من الجيد أن يستند هذا النوع من الالتزامات إلى أساس قانوني من نوع ما، كأن يدخل مثلاً في تشريع يمنحه صلاحية. وهكذا يمكن اتخاذ إجراءات قضائية بحق مرتكبي المخالفات. لكن هذا لا يمنع اعتماد إجراءات أخرى تحكم الانتهاكات، مثل رفض تقديم خدمات أخرى إلى الباحث المعني و/أو في نهاية المطاف المؤسسة التي ينتمي إليها“ [24]

إذا كان مكتب الإحصاءات الوطني يرغب في الحصول على آراء المستخدمين، ينبغي أن يضع تدابير متابعة منتظمة للباحثين. ويمكن اقتناص المناسبة لتذكيرهم بواجبات إيصال نتائجهم وسؤالهم عن اقتراحاتهم لتحسين برنامج الاستقصاء.

الإطار 12 شروط النفاذ إلى ملفات البيانات الخاضعة لترخيص واستخدامها

ملاحظة: إن النقاط 1 إلى 8 هي نفسها الشروط المطبقة على الملفات المتاحة للاستخدام العام. أما النقطتان 9 و10 فيجب تكييفها لتلائم العقد الإطار.

1. لا يمكن بأي حال من الأحوال بيع البيانات والوثائق الأخرى التي يقدمها مكتب الإحصاءات الوطني من جديد أو إعطاؤها إلى أشخاص، أو مؤسسات، أو منظمات أخرى من دون موافقة خطية مسبقة من طرف مكتب الإحصاءات الوطني.

2. تستخدم البيانات حصراً لغايات إحصائية أو في إطار أعمال بحث علمي. كما أنها تهدف فقط إلى تشكيل بيانات مجمعة، بما في ذلك تصميم المعلومات. أما البحث عن أشخاص أو منظمات بحد ذاتها فمحظور بجميع أشكاله.

3. تحظر أي محاولة لإعادة تحديد المستجيبين. وعلاوة على ذلك، يمنع استخدام هوية أي شخص أو منظمة تم اكتشافها مصادفةً. وينبغي إبلاغ مركز الإحصاءات الوطني بأي اكتشاف من هذا النوع.

4. تحظر أي محاولة لإعادة اقتطاع مجموع البيانات التي أتاحها مكتب الإحصاءات الوطني، سواء مع بيانات مكتب الإحصاءات الوطني أو غيرها من مجموعات البيانات التي من شأنها أن تسمح بالتعرّف إلى هوية أشخاص أو منظمات.

5. يجب أن تحتوي كل الكتب أو المقالات أو وثائق العمل أو الدراسات أو الأطروحات أو التقارير أو غيرها من المنشورات المستندة إلى البيانات التي قدّمها مكتب الإحصاءات الوطني على مراجع لمصدرها، وفقاً لطلب ذكر مراجع الاقتباسات الخاص بمجموع البيانات المقدّمة.

6. تقدّم إلى مكتب الإحصاءات الوطني نسخة بصيغة إلكترونية تضم كل المنشورات المستندة إلى البيانات المطلوبة.

7. لا تتحمل الجهة التي جمعت البيانات في المرحلة الأولى، ولا مكتب الإحصاءات الوطني، ولا الجهات الممولة المعنية بأي شكل من الأشكال المسؤولية المرتبطة باستخدام هذه البيانات، أو تحليلها، أو النتائج المستخلصة منها.

8. يجب ذكر اسم المؤسسة المعنية، واسم الباحث والباحثين الآخرين الذين سيستخدمون البيانات. ويجب أن يوقع الباحث الرئيسي عقد ترخيص باسم المؤسسة. وإذا لم يكن مخولاً للتوقيع باسم المؤسسة المتلقية، ينبغي تحديد اسم ممثل الأخيرة.

9. يجب تحديد غايات الاستخدام المرجوة من البيانات، مع قائمة بالبيانات المتوقعة وسياسة النشر التي تعتمدها المؤسسة.

(يمكن للشريطين 8 و9 أن يتغيّرا في حالة مؤسسات التعليم العالي)

الإطار 13 العقد الإطاري

9. تلتزم [الجهة المستفيدة] بأن تقدّم إلى [الجهة المانحة] تقريراً سنوياً عن استخدام مجموعات البيانات المذكورة أعلاه ومستخدميها، محددة عدد الباحثين المتمتعين بحق النفاذ إلى كل مجموعة من البيانات، ونتائج أعمال البحث التي تجرى.
10. يمنح حق النفاذ إلى البيانات لفترة (فترة محددة أو عقد لأجل غير محدد).

ج. التغييرات:

تمثل [الجهة المستفيدة] محادناً واحداً في هذا العقد. وإذا تم تغيير هذا الطرف، تتعهد [الجهة المستفيدة] بأن تقدّم إلى [الجهة المانحة] من دون تأخير اسم المحادث الجديد ومعلومات الاتصال به. ويمكن أن يتم الإبلاغ عن أي تغيير إداري أو تغيير متعلق بقاعدة معبنة عن طريق البريد الإلكتروني، أو الفاكس، أو البريد وتحديد المعلومات التالية:

التغييرات التي تعلم [الجهة المانحة] [الجهة المستفيدة] بها:

اسم الشخص الذي يجدر الاتصال به: _____
وظيفة الشخص الذي يجدر الاتصال به: _____
عنوان المتلقي: _____
البريد الإلكتروني: _____
الفاكس: _____

التغييرات التي تعلم [الجهة المستفيدة] [الجهة المودعة] بها:

اسم الشخص الذي يجدر الاتصال به: _____
وظيفة الشخص الذي يجدر الاتصال به: _____
عنوان المتلقي: _____
البريد الإلكتروني: _____
الفاكس: _____

د. الطرفان الموقعان

قرأ الطرفان الموقعان أدناه أحكام هذا العقد وقبلها بها:

ممثل [الجهة المانحة] _____
الاسم _____
التوقيع _____ التاريخ _____

ممثل [الجهة المستفيدة]

الاسم _____
التوقيع _____ التاريخ _____

عقد بين [الجهة المانحة] و[الجهة المستفيدة] فيما يتعلق بتقديم البيانات الجزئية واستخدامها
أ. يعنى هذا العقد مجموعات البيانات الجزئية التالية:

- _____
- _____
- _____
- _____
- _____

ب. أحكام العقد:

إن ممثل [الجهة المانحة]، بصفته صاحب حقوق الملكية الخاصة بالمنتجات المذكورة في القسم أ، أو بموجب الترخيص الذي منحه إليه الأخيرة بطريقة سليمة ورسمية، يقبل بوضع مجموعات البيانات المذكورة في القسم أ في تصرف [الجهة المستفيدة]، على أن يحصر استخدامها لموظفي [الجهة المستفيدة]، وبشرط احترام الشروط التالية:

- لا يمكن بأي حال من الأحوال بيع البيانات (بما فيها المجموعات الفرعية للبيانات) وغيرها من العناصر التي تحميها حقوق الملكية والتي قدّمها [الجهة المانحة] من جديد أو إعطاؤها إلى أشخاص، أو مؤسسات، أو منظمات أخرى من دون موافقة خطية مسبقة من [الجهة المانحة]. أما العناصر التي لا تحميها حقوق الملكية والتي لا تحتوي على بيانات جزئية (مثل نماذج استطلاعات الرأي، والإرشادات، ومعاجم القوانين أو البيانات) فيمكن إيصالها بالمقابل من دون موافقتها المسبقة. وتبقى [الجهة المانحة] مالكة لمجموع المنتجات التي أتيحت تحت رعايتها.
- تستخدم البيانات حصراً لغايات إحصائية أو في إطار أعمال بحث علمي. كما أنها تهدف فقط إلى تشكيل بيانات مجمعة، بما في ذلك تصميم المعلومات. أما البحث عن أشخاص أو منظمات بحد ذاتها فمحظور بجميع أشكاله.
- تحظر أي محاولة لإعادة تحديد المستجيبين. وعلاوة على ذلك، يمنع استخدام هوية أي شخص أو منظمة تم اكتشافها مصادفةً. وينبغي إبلاغ [الجهة المانحة] بأي اكتشاف من هذا النوع.
- تحظر أي محاولة لإعادة اقتطاع مجموع البيانات التي أتاحتها [الجهة المانحة]، سواء مع بيانات [الجهة المانحة] أو غيرها من مجموعات البيانات التي من شأنها أن تسمح بالتعرّف إلى هوية أشخاص أو منظمات.
- يجب أن تحتوي كل الكتب، أو المقالات، أو وثائق العمل، أو الدراسات، أو الأطروحات، أو التقارير، أو غيرها من المنشورات المستندة إلى البيانات التي قدّمها مكتب الإحصاءات الوطني على مراجع لمصدرها، وفقاً لطلب ذكر مراجع الاقتباسات الخاص بمجموع البيانات المقدمّة.
- تقدّم إلى [الجهة المانحة] نسخة بصيغة إلكترونية تضم كل المنشورات المستندة إلى البيانات المطلوبة.
- لا تتحمل [الجهة المانحة] ولا الجهات الممولة المعنية بأي شكل من الأشكال المسؤولية المرتبطة باستخدام هذه البيانات، أو تحليلها، أو النتائج المستخلصة منها.
- تحفظ البيانات في بيئة آمنة ويقيد النفاذ إليها بطريقة ملائمة. وتحفظ [الجهة المانحة] بحق طلب معلومات في أي وقت كان عن تدابير التخزين وعن أنظمة النشر المستخدمة.

7. ماذا نعني بـ“إغفال” البيانات الجزئية؟

في معظم البلدان النامية، يعدّ خطر الكشف المرتبط بإمكانية إعادة اقتطاع ملف بيانات جزئية لاستقصاء ما مع ملفات بيانات أخرى محدوداً حالياً، سواء لأنها غير موجودة بعد أو لأنها ليست منشورة على نطاق واسع بما فيه الكفاية. وفي الحقيقة، يمكن تقليص خطر الكشف المرتبط بنشر ملفات البيانات الجزئية تقليصاً مُرضياً، وذلك ببساطة عن طريق حذف المعرّفات المباشرة للتسجيلات، والمعلومات الجغرافية الفائقة الدقة مقارنة بنطاق أخذ العينات. ومع ذلك، يجدر تقسيم خطر الكشف لكل حالة على حدى. وفي الواقع، ثمة استثناءات جديرة بالذكر للقاعدة العامة المذكورة أعلاه، ما يفرض إجراءات الضبط الإضافية، أما إغفال ملفات البيانات الجزئية الخاصة باستقصاء يحتوي على تسجيلات متعلقة بمجموعات مستهدفة ضيقة، مثل مؤسسة على سبيل المثال، فهو أصعب بكثير.

وأخيراً، على مكاتب الإحصاءات الوطنية أن تأخذ أمراً آخر في الاعتبار: ألا وهو أهمية حماية قواعد عيناتها، لا سيما قاعدة العينات الرئيسية المتداخلة بين الاستقصاءات، والمستعملة لتصميم استقصاءات أخرى لدى الأسر. أما نشر قواعد أخذ العينات المفصلة فهو لا يتلاءم مع الإجراءات الرامية إلى حماية هوية المستجيبين. ويمكن لهذه المعلومات غير المرصّدة أن تستخدم لإنشاء مقاطع، قد تتيح بالتالي التعرف إلى معلومات اتصال جغرافية (اسم وحدة العينة الأولية ومكانها) عن طريق اقتطاع عناصر الترحيب والعيّنات (التي يجب أن تظهر في ملفات البيانات الجزئية حتى يتمكن المستخدمون من التوصل إلى استنتاجات إحصائية صحيحة). إن حماية قواعد العينات تُفسّر بالرغبة في تقليص احتمال اختيار المستجيبين المنتمين إلى وحدة عينات أولية ليشاركوا في استقصاءات أخرى لا يجريها مكتب الإحصاءات الوطني. وفي البلدان التي لا يتم فيها تحديث وحدات العينات الأولية كل عشر سنوات، قد ينتج عن ذلك ارتفاع في نسبة عدم الاستجابة (ويشار إليهم بالمستجيبين السُتمين) وقد يتسبب أيضاً في بعض الالتواءات.

7.2 سيناريوهات الكشف

تتمثل الخطوة الأولى في عملية إغفال ملف بيانات معدّ للنشر في تحديد أجزاء الملف التي قد تُسهّل الكشف. وبالتالي، ينبغي قبل وضع إجراءات ضبط الكشف عن الإحصاءات فحص الحالات المختلفة التي قد يستطيع فيها مستخدم ملف بيانات جزئية تحديد هوية أحد المستجيبين. وتميّز في هذا الإطار بين حالتين:

- **معرفة الإجابة:** يتمتع المستخدم، بصفته الشخصية، بما يكفي من المعلومات عن الملحقات المتعلقة بأحد التسجيلات أو عدد منها. وبعبارة أخرى، ينتمي إلى دائرة معارف وحدة إحصائية. وكلما كانت المجموعة المستهدفة محصورة (استقصاءات في قلب شركة أو مؤسسة، أو استقصاءات لدى الأسر في بلد صغير أو مناطق ذات كثافة سكانية منخفضة مثلاً).
- **إعادة الاقتطاع مع أرشيفات خارجية:** يجمع المستخدم بعض التسجيلات من ملف البيانات الجزئية المنشور إلى جانب غيرها من مجموعات البيانات (أو السجلات) العامة التي تحتوي على معرّفات مباشرة، على الرغم من الحظر المنصوص عليه في عقد استخدام البيانات. ولهذه الغاية، يستخدم متغيرات مفتاح متاحة في مجموعتي البيانات ليشكل مفاتيح تعريف (مقارنة البيانات).

قبل إغفال البيانات، ينبغي تصوّر سيناريو كشف يحدد المعلومات التي قد تنشر، والطرق المحتملة للتعرف إلى هوية الأشخاص. ويستند هذا السيناريو عادة

يستند هذا الفصل إلى حد كبير على الدليل الذي نشرته CENEX بشأن ضبط الكشف عن الإحصاءات (3)، الذي يمكن تحميله مجاناً على الرابط التالي: http://neon.vb.cbs.nl/case/SDC_Handbook.pdf.
(متوفر بالإنجليزية فقط)⁹

تعدّ مهمة نشر ملفات البيانات الجزئية إلى العامة، أو الباحثين أو مؤسسات أخرى مهمة دقيقة بالنسبة إلى مكاتب الإحصاء الوطنية. إذ تتمثل من جهة في إتاحة ملفات البيانات الجزئية بالاستناد إلى عدد كبير من التحاليل الإحصائية، ومن جهة أخرى في الحفاظ على هوية المستجيبين. أما العملية المتعلقة بالجانب الثاني، فيشار إليها عادة بمفهوم ضبط الكشف عن الإحصاءات (SDC) أو الإغفال.

”لا يمكن ضمان السرية بنسبة 100%، لذا يجب قياس المخاطر والمصالح المرتبطة بالنفاذ إلى البيانات وفقاً لمخاطر الكشف عن البيانات وحساسيتها“ [14]. وبالنسبة إلى مكاتب الإحصاءات وغيرها من منتجي البيانات، يعدّ السؤال عن حماية البيانات سؤالاً مشروعاً تماماً، إذ تحيل الدراسات التي أجريت في مجال إدارة سرية البيانات هي الأخرى في مرا كثيرة إلى ما أسماه بيتر مادسن Peter Madsen (13) ”تناقض السرية“. ففي عام 2003، وفي ورشة عمل مكرسة لدراسة السرية نظمت برعاية National Science Foundation الأمريكية، أشار بيتر مادسن إلى أن ”الحساس الذي يكرّس بهدف حماية البيانات بالكامل في إطار البحث يؤدي، وبشكل معاكس، إلى تقليص الفائدة بالنسبة إلى المجتمع عوضاً عن رفعها“¹⁰.

7.1 مفاهيم مرتبطة بضبط الكشف عن الإحصاءات

يحدث الكشف حين يلاحظ شخص أو مؤسسة أو يطلع على معلومة ممنوعة عن شخص أو مؤسسة أخرى. وتميّز بين نوعين من مخاطر الكشف: هما **الكشف عن الهوية والكشف عن الملحقات**¹¹. ويتم الكشف عن الهوية حين تربط هذه الأخيرة مباشرة بتسجيل في ملف البيانات المنشور. ويحدث ذلك عندما يضم التسجيل المعني متغيرات تسمح بالتعرّف إلى مستجيب ما بوضوح (اسمه، أو رقم جواز سفره/ رقم تعريفه الوطني، أو معلومات الاتصال الهاتفي به مثلاً). ومن الضروري حذف هذه المتغيرات من كل ملف بيانات جزئية قبل نشره. أما الكشف عن الملحقات، فيحدث حين يمكن ربط القيمة (الدقيقة أو المقدّرة) الواردة في البيانات المنشورة بشخص بالذات.

وتطلق تسمية ”المفتاح“ (أو المتغيرة المفتاح) على مجموعة المتغيرات الخاصة بتسجيل ما في البيانات الجزئية، والتي تتيح إمكانية الوصول إلى المستجيب. إن ”إعادة التعريف“ هذه ممكنة حين (أ) ينتمي مستجيب إلى فئة من السكان تظهر قيمة مفتاح معينة نادرة؛ و (ب) حين يتيح المفتاح إعادة اقتطاع ملف من البيانات الجزئية مع غيره من ملفات البيانات التي قد تحتوي على معرّفات مباشرة أو غيرها (قوائم المقترعين، أو السجلات العقارية، أو الملفات المدرسية، أو حتى باستخدام محرّكات البحث العامة على شبكة الإنترنت).

⁹ يحتوي الموقع الإلكتروني لـ American Statistical Society، الذي يحدث دائماً، أيضاً على معلومات ووثائق عن ضبط الكشف عن الإحصاءات. أنظر <http://www.amstat.org/committees/cmtepc/index.cfm?fuseaction=main> (متوفر بالإنجليزية فقط).

¹⁰ يقدّم الرابط التالي خلاصة نتائج ورشة العمل هذه: http://www.nsf.gov/sbe/ses/mms/nsfworkshop_summary1.pdf. (باللغة الإنجليزية فقط)

¹¹ اقترح المصطلحات اداكتور لامبرت (12).

يبلغ من العمر 30 سنة، يمارس مهنة طبيب، أب لأربعة أطفال إناث، الوحيد الذي يملك هذه المواصفات في عينة الاستقصاء. عندئذٍ يعتبر التسجيل حساساً، حتى وإن كان السكان الإجماليون يضمنون أفراداً آخرين يتمتعون بالميزات نفسها. كما يمكن تحديد الخطر الذي تواجهه وحدة ما عن طريق جمع تعداد متغيرات التعريفات النوعية في قلب مجموع السكان أو عن طريق ترجيح إعادة التعرّف إليه. أما فرصة التكرار في قلب مجموع السكان فعادة ما يكون مجهولاً، لذا يجب إجراء هذه الترحيحات عن طريق تصميم. ومع متغيرات تعريفية متصلة، يستحيل استغلال سمة ندرة المفاتيح، التي تكون في معظمها، إن لم تكن كلها، فريدة. ويقدّر خطر الكشف عن طريق المتغيرات المتصلة وفقاً لتقدير احتمال إعادة التعرّف بواسطة مقارنة متغيرات مجموعتي بيانات منفصلة، على قاعدة "تجاوز" قيمها.

يمكن التوصل إلى الكشف عن المعلومات بواسطة الكثير من طرق التعرّف إلى الحالات والمتغيرات. لذا تتمثل إحدى الطرق الشائعة في إجراء توزيعات تكرارية وجدول متعددة الأبعاد، من أجل تحديد الخلايا التي تعود إلى عدد قليل من الحالات. كما تندرج المعلومات الجغرافية المفضلة في أبرز مصادر التعرّف على الأفراد، لا سيما حين يتناولها مستخدمون ينتمون إلى المنطقة عينها ويعرفون جيداً سمات بعض المستجيبين.

الإطار 14 قائمة النقاط التي ينبغي التحقق منها في إطار تقييم مختلف الحالات ومختلف مخاطر الكشف

في إطار الفحص الذي يسبق نشر ملفات البيانات الجزئية العائدة إلى الإحصاءات والمسوح التي يتكلف مكتب Census Bureau الأمريكي بحماية سرية مستجيبها، ثمة قائمة بالنقاط التي ينبغي التحقق منها (www.census.gov/srd/sdc/) on Disclosure Potential of Proposed Data Releases. Checklist). تسهّل هذه القائمة عملية ضبط الشف عن السرية. وهي تهدف إلى مساعدة الأشخاص المكلفين بإجراء هذا الضبط في تحديد ما إن كان يمكن إتاحة ملفات البيانات الجزئية أو البيانات المجدولة للاستخدام العام. ويتناول القسم الثالث من هذه القائمة ملفات البيانات الجزئية. أما النقاط التي ينبغي التحقق منها، فتتعلق بعناصر مهمة مثل معلومات الاتصال الجغرافية، والمتغيرات التي قد تتفاقم خطر الكشف بشكل غير اعتيادي، والمعلومات السياقية والبيئية، وحمية تقارب أو مقارنة مع بيانات أخرى، وعمليات إعادة الاقتطاع المحتملة التي يمكن أن تؤدي إلى تعرّف على مجموعة فريدة من الملحقات.

7.4 تقنيات ضبط الكشف عن السرية الخاصة بملفات البيانات الجزئية

تتمثل الخطوة الكبرى الأولى من عملية ضبط الكشف عن سرية ملفات البيانات الجزئية في حذف كل المعرفات المباشرة، أي المتغيرات التي تسمح بالتعرّف إلى مستجيب ما بطريقة أحادية. يلي ذلك عملية إغفال ملف البيانات الجزئية. وهي تستعين بطرق "الإخفاء" أو تنشئ لهذه الغاية بيانات جزئية مركبة. وينشأ ملف البيانات الجزئية المركبة بطريقة عشوائية عن طريق عملية تسمح بالحفاظ على بعض التلازمات الإحصائية أو الداخلية في ملف البيانات الجزئية الصافي الأصلي.

ويعنى بطرق الإخفاء التقنيات التي تسمح بإنشاء نسخة معدلة من ملف البيانات الجزئية الصافي الأصلي. وتمييز بين نوعين من وسائل ضبط الكشف عن السرية عن طريق الإخفاء. أولها الوسائل المستندة إلى تشويش المعلومات، وهي تتمثل في تعديل البيانات قبل نشرها عن طريق إدخال عنصر خاطئ فيها عن قصد

إلى فرضيات حذرة، تتوقع أسوأ الحالات الممكنة. وأحياناً، من الضروري تحديد سيناريوهات عدة، إذ قد يتم اقتراح مصادر معلومات مختلفة على المستخدم، سواء في وقت واحد أو بالتناوب.

7.3 تقييم خطر الكشف

تنتهك السرية في حال التعرف إلى أحد المستجيبين وحين يتمكن المنتهك (المستخدم غير المرخص أو الذي لا يحترم الشروط الواردة في عقد استخدام البيانات والنفاد إليها) من ربط المتغيرات الحساسة بأحد الأفراد.

قبل نشر البيانات الجزئية، يتعيّن على مكاتب الإحصاء الوطنية أن تقيّم البيانات "لمعرفة ما إذا كان نشرها يضع على المحك هوية بعض الأشخاص أو المؤسسات. وتؤخذ عوامل عدة في الاعتبار في هذا الإطار: (1) مستوى التفاصيل الواردة في هذه البيانات المعدّة للنشر (لا سيما التفاصيل الجغرافية والمتغيرات التي تعدّ شائعة في مصادر البيانات الخارجية، التي من شأنها أن تمثل مفاتيح وتزيد من خطر التعرّف إلى الأشخاص)؛ (2) بعض المتغيرات أو مجموعات المتغيرات التي تسمح بعزل مستجيبين في قلب العينة والتي قد تزيد من احتمال تعرّف أشخاص خارجيين عليهم؛ و(3) وجود بيانات أخرى متوافرة خارج (مكتب الإحصاءات الوطني)، مثل تلك التي سبق ونشرت عن الاستقصاء عينه أو عن استقصاء مواز، أو حتى معلومات تملكها مؤسسات أخرى عن المستجيب عينه". [14]

"تشكل مهمة تحديد خطر الكشف عن البيانات التي قد تتيح التعرّف إلى المستجيبين مهمة معقدة، وهي تستلزم تحليلاً إحصائياً تجريبياً وفطنة في أن معاً" [14]. وثمة طرق كثيرة تمثل كلها ميزات يمكن استغلالها. ولكن لا تفرض أي طريقة منها على اعتبارها الطريقة "الفضلى". وبالطبع، لا يمكن استبعاد خطر الكشف استبعاداً نهائياً، ولكن ثمة إجراءات يمكن اتخاذها عن طريق تقليص نفعية البيانات. وتشمل هذه الطرق وضع عتبة للخطر (threshold rule) تسمح بمعرفة ما إذا كان نشر مجموعة البيانات آمناً أم لا. وثمة طريقتان رئيسيتان لقياس خطر إعادة التعرف إلى المستجيبين من منطلق رياضي:

- **الإجراءات الفردية:** تقييم الخطر لكل تسجيل. وترد هذه الإجراءات عادة بشكل ترجيحات لإعادة التعرّف إلى مستجيب من حيث التمييز والندرة في قلب العينة.
- **الإجراءات العامة:** تقييم الخطر بالنسبة إلى مجموع الملف. وتوصل هذه الإجراءات إلى عدد تقريبي لاحتمالات إعادة التعرّف ويمكن اشتقاقها من الإجراءات الفردية المجمعة.

تكمّن فائدة الإجراءات الفردية في أن التسجيلات التي تبدو حساسة من وجهة نظر عتبة الخطر هي وحدها التي يتعيّن حمايتها. ويحدّ ذلك من خسارة معلومات البيانات ونفعيتها. وقد يشجع الاكتفاء بالإجراءات العامة على اللجوء إلى تقنيات ضبط الكشف عن السرية لكل تسجيل في ملف البيانات الجزئية، ما يزيد من خطر خسارة المعلومات ويقلص فوائد البيانات للتحليل الإحصائي.

يمكن تمييز طرق قياس مخاطر إعادة التعرّف أيضاً عن طريق استخدام المفاتيح. إذ تسمح القياسات المستندة إلى مفاتيح العينة بالتعرّف إلى مجموعات المتغيرات الكمية (المفاتيح) الفريدة أو النادرة في قلب عينة ما. وتعدّ الوحدة حساسة حين يكون مجموع نقاط متغيرات التعريف الخاصة بها أدنى من العتبة المتفق عليها. فعلى سبيل المثال، من المرجح للغاية أن يكون شخص من نوع ذكر،

لغايات حماية السرية، أما ثانياً فهي الوسائل المستندة إلى حظر المعلومات، وهي تقتضي تقليص المحتوى المعلوماتي في البيانات المقدّمة عن طريق حذف جزء منها أو عن طريق تجميعها.

فيما يلي مقدمة عامة تتناول التقنيات الأكثر شيوعاً بضبط الكشف عن السرية، مع ذكر نوع البيانات التي تطبق عليها: المتغيرات المتصلة أو النوعية، أو النوعين معاً، ونقول عن متغيرة ما إنها متصلة إذا كانت رقمية وتتيح إجراء عمليات حسابية (مثل المداحيل، والأعمار، وعدد أفراد الأسرة). أما المتغيرات التي تحدد فقط لمجموعة نهائية ولا يمكن إجراء حسابات رياضية من خلالها، فنسميها المتغيرات النوعية أو الفئوية (المتغيرات المدرجة على سلم ترتيبي، مثل أعلى مستوى محقق لارتياح المدارس، أو المتغيرات التي تقاس بسلم اسمي، مثل الوضع العائلي، حيث يكون ترتيب القيم غير مهم).

التقنيات المستندة إلى حظر المعلومات

تقضي التقنيات المستندة إلى حظر البيانات المقدّمة أو محتواها المعلوماتي بتعديل ملفات البيانات الجزئية بهدف إبعاد المتغيرات أو التسجيلات التي قد تربط بأحد الأفراد دون غيره من دون أي لبس. أما الخيار الآخر فيتمثل في إنشاء فئات، بحيث يزيد عدد المستجيبين المحتملين لهذه الفئة المعينة، وهكذا يستطيع مكتب الإحصاءات الوطني أن يحدد عتبة تفرض وجود حد أدنى من الإجابات للخلية الواحدة، وفي ما يلي تقنيات الإخفاء الست الكبرى المستندة إلى حظر المعلومات:

1. **أخذ عينات:** بدلاً من نشر ملف البيانات الجزئية بكامله؛ وهي طريقة يوصى بها لنشر بيانات مسوح السكان. وتسمح هذه التقنية بتقليص خطر الكشف بقدر كاف للبيانات الكمية، شرط أن تحصر العينة في نطاق محدود بما فيه الكفاية (5% من السكان مثلاً) وشرط حذف كل متغيرات التعرّف المباشرة. ومع ذلك، إذا كان ملف البيانات الجزئية يحتوي على متغيرات متصلة، قد يتم اقتطاعها بسهولة أكبر مع ملف بيانات متاح في مكان آخر، فيتوجب عندئذ تنفيذ تقنيات أخرى من تقنيات ضبط الكشف عن السرية.
2. **إعادة الترميز الشاملة:** تقتضي بإعادة تجميع قيم وفقاً لتصنيفات محددة مسبقاً، بحيث لا تظهر الإجابات الفردية. وتطبق هذه الطريقة على المتغيرات المتصلة أو المنفصلة بقدر ما تطبق على الرموز الجغرافية. فعلى سبيل المثال، قد يقطع العمر ويستبدل به فئات عمرية، كما قد تجمع المهنة والرموز الصناعية في فئات أوسع. بالإضافة إلى ذلك، يمكن حذف المعلومات الجغرافية الأكثر دقة من تلك التي تحدد نطاق أخذ العينات، وتعتمد طريقة إعادة الترميز الشاملة للبيانات المتصلة والنوعية على السواء.
3. **إعادة تجميع القيم العليا والدنيا:** يتم اللجوء إلى هذه العملية في حال وجود متغيرات رقمية أو ترتيبية تندر فيها القيم العليا والدنيا المتطرفة، الأمر الذي يمكن أن يتيح التعرّف إلى هوية المستجيبين. وتقتضي إعادة تجميع القيم المتطرفة العليا والدنيا إنشاء فئات "خشنة" (للقيم العليا، والأشخاص الذين يفوق عمرهم الـ X سنة، أو الدخل الأعلى من Y على سبيل المثال). وتطبق هذه التقنية على المتغيرات المتصلة وتلك النوعية التي تقاس وفقاً لسلم ترتيبي، والتي يمكن أن يحدد لها منطقتين لتنظيمها في فئات.
4. **الحذف الموضوعي:** هي تقنية بسيطة للغاية تستخدم حين يمكن أن يؤدي تقارب متغيرتين فرديتين إلى التعرّف على شخص معين.

وبعبارة أخرى، يمكن لمجموعة من المتغيرات أن تعطي مفتاحاً لإعادة التعرّف في تسجيل ما. لنأخذ مثلاً تسجيل ما. لنأخذ مثلاً تسجيل ما. لتالية: العمر = 58، والوضع الدراسي = مسجل حالياً في المدرسة الابتدائية. يرجح أن يكون هذا التسجيل منقطع النظر، إذ نادراً ما يسجل الأشخاص في عقدهم التاسع في المدرسة الابتدائية¹². في هذه الحالة بالذات، يكفي حتماً حذف إحدى هاتين المعلوماتين (العمر أو الوضع الدراسي) للتخلص من المشكلة. ويعدّ الحذف الموضوعي مفيداً لا سيما في البيانات النوعية.

5. حذف بعض المتغيرات: تُطلّب هذه الطريقة عند التعامل مع معلومات تعدّ غير ملائمة للنشر لأنها حساسة للغاية (الانتماء العرقي أو الدين مثلاً).

6. حذف بعض التسجيلات: عملية مهمة أحياناً للمحافظة على مجهولية هوية المستجيبين مع مجموعة واحدة من المتغيرات. فحين يحذف تسجيل بالكامل من ملف البيانات الجزئية، يصبح من الإلزامي اللجوء إلى حساب عوامل ترجيح معدّلة وتوقعها. كما يتعيّن استخدام هذه الطريقة عند اللزوم فقط، لأن حذف عدد كبير من التسجيلات يؤدي إلى تشويش في البيانات.

تقنيات التشويش

تقضي تقنيات التشويش بتعديل البيانات لجعلها أقل دقة وجعل مهمة إعادة اقتطاعها أكثر صعوبة. وفي حال محاولة إعادة التعرّف، فإن القيم المغيّرة تجعل الترابطات غير أكيدة، وفيما يلي موجز يعرّف بتقنيات التشويش الكبرى:

1. **إضافة الضجيج (أو التشويش العشوائي):** هي تقنية تقتضي إضافة قيم عشوائية إلى تلك التي نقلت عن المستجيب، وتختلف الطرق التي تهدف إلى إضافة ضجيج إلى إحدى المتغيرات أو عدد منها من دون أن يمس ذلك المعدلات، والفروق والتغايرات. كما تسمح تقنيات البرمجة الخطية بتقليص الاختلافات بين القيم المعدّلة والقيم الحقيقية.
2. **تبديل البيانات:** هي طريقة تقتضي تعديل ملف بيانات جزئية بحذف قيم المتغيرات السرية ووضع قيم تسجيلات أخرى مكانها، وهكذا "يتم تبديل" التسجيلات بحيث تحفظ القيم الهامشية أو الخلايا العائدة إلى عدد من الحالات. ويمكن استخدام هذه التقنية للمتغيرات المتصلة والنوعية على السواء.
3. **تبديل الرتبة:** هو نظام يسمح بفرز المتغيرات المتوجب حمايتها بالترتيب التصاعدي، ومن ثم إعادة تجميع بعض منها. وتنتقى أزواج من التسجيلات بطريقة عشوائية في كل مجموعة، وتبدّل قيمها بقيم أزواج أخرى في نطاق محدد مسبقاً. ويؤدي إنشاء مجموعات مختلفة الأحجام إلى الحصول على نماذج مختلفة من البيانات.
4. **التجميع الجزئي:** يقتضي تغيير قيمة تمت ملاحظتها في العينة ليوضع مكانها المعدّل المحسوب لمجموعة محصورة من الوحدات (مجموع صغير، أو مجموع جزئي)، من بينها الوحدة المضبوطة، وتمثل الوحدات المنتمية إلى المجموعة نفسها بالقيمة نفسها في الملف المنشور. تحتوي كل مجموعة على حد أدنى محدد مسبقاً من الوحدات (k). أما القيمة الدنيا المقبولة لـ k فهي 3. فبالنسبة إلى قيمة من k، يتمثل الهدف في تقسيم مجموع الوحدات في مجموعات فيها على الأقل k وحدات، وذلك بتقليص خسارة المعلومات إلى الحد الأدنى،

الباحثين المعتادين إجراء تحليل معمّقة مستندة إلى ملفات البيانات الجزئية. وغالباً ما يكون هؤلاء قلة من الباحثين الذين تصبّ أعمالهم في المصلحة العام إلى حد بعيد. ولعل ذلك يظهر أهمية عدم الاكتفاء ببث ملفات البيانات الجزئية المتاحة للاستخدام العام، بل إرفاقها بالملفات المرخصة الأقل إغفالاً. ومن ناحية أخرى، يجدر بمكاتب الإحصاءات الوطنية أن تكون فكرة واضحة عن خسارة المعلومات الناتجة عن إجراءات ضبط الكشف عن الإحصاءات، وذلك عن طريق فحص استمارات طلبات النفاذ إلى الملفات الخاضعة لترخيص، التي تذكر الأسباب التي تجعل النسخة المتاحة للاستخدام العام غير مفيدة في إطار مشاريع الأبحاث المعنية.

7.6. توثيق عملية ضبط الكشف عن البيانات الإحصائية

يجب أن توازن طرق ضبط الكشف عن الإحصاءات بين خسارة المعلومات واحتمال الكشف عن المعلومات الشخصية. إذ يتعيّن إبلاغ المستخدمين عند الاقتضاء بتقييم خطر الكشف عن مجموع البيانات، وبطرق الحماية المطبقة. كما يفضل أن يكونوا على اطلاع على طبيعة كل تعديل ونطاقه في هذا الصدد.

ويمكن تحديد طريقة (أو طرق) لضبط الكشف عن البيانات، غير أن مستوى التفاصيل المتوافرة يجب ألا أن يسمح للمستخدم بإعادة بناء ملفات البيانات الجزئية الأصلية بأي شكل من الأشكال.

الأمر الذي ينعكس عادة في خسارة للتغاير. يعاد توزيع الوحدات إذاً في مجموعات وفقاً لعدد التشابهات بينها. وهكذا تحمي عملية التجميع الجزئي البيانات، إذ تجعل الملف يحتوي على الأقل k وحدات مجمّعة بالقيمة نفسها، وتطبّق هذه التقنية أحياناً على المتغيرات المتصلة.

5. التدوير: تستخدم وسائل متعددة للتدوير وفقاً للهدف المنشود: فثمة التدوير المضبوط ليحفظ الأعداد الإجمالية وبعض الخصائص العامة، وثمة التدوير العشوائي بحيث لا يكشف عدد الحالة الظاهر في الخلية في جدول البيانات المجمّعة عن نتائج عينة أو عينتين.

6. إعادة أخذ العينات: تقتضي هذه العملية اختيار عينات مختلفة مستقلة عن قيم المتغيرات المعدة لتخفي. ويبقى معيار التصنيف هو نفسه. وفي هذا السياق، تنشأ المتغيرات المخفية كما يلي: تكون القيمة الأولى معدّل جميع القيم الأولى لكل عينة، إلخ.

7. إضفاء العشوائية في مرحلة متأخرة: نسخة عشوائية لتبديل البيانات. تضيف هذه التقنية عنصراً غير أكيداً إلى القيم العائدة إلى بعض المتغيرات عن طريق تغييرها وفقاً لآلية ترجيحية. وكما في حالة التبديل، تتحقق حماية البيانات، لأن المستخدمين لا يعلمون يقيناً دقة القيم التي حصلوا عليها. وبالتالي، فإن أي محاولة لاقطاع التسجيل مع معرّفات خارجية، من شأنها أن تؤدي إلى أخطاء في التصنيف أو التحليل. ويتم اللجوء إلى هذه التقنية بشكل رئيسي في حالة المتغيرات النوعية، ولكن يمكن تطبيقها أيضاً على المتغيرات الرقمية المتصلة.

ملفات البيانات الجزئية المركبة

تنشأ ملفات البيانات الجزئية المركبة على أساس عشوائي، أما القاعدة التي ينبغي احترامها في هذا الصدد فتتمثل في المحافظة على الارتباطات الإحصائية أو الداخلية في الملف الأصلي. لا شك إنه لمن المعري تقديم ملفات مؤلفة بالكامل بواسطة تناظرات ولا تمثل أي خطر كشف عوضاً عن تقديم ملفات البيانات الجزئية الأصلية. لذا وضعت لهذه الغاية تقنيات متعددة. ولكن في مجال طرق إخفاء البيانات، تعدّ مختلف التقنيات الرامية إلى إنشاء ملفات بيانات جزئية مركبة، كلها تقنيات فائقة التعقيد والصعوبة. إذ تضم بيانات الاستقصاء عادة مئات المتغيرات، أما توزيعها وارتباطاتها فليست سهلة التنظيم بواسطة أدوات إعداد معيارية. وما زالت الأبحاث الرامية إلى تحسين تطبيق هذه التقنيات جارية.

7.5. إيجاد الحل الوسط بين خطر الكشف وخسارة المعلومات

يؤدي تغيير ملف البيانات الجزئية عن طريق تقنيات ضبط الكشف عن الإحصاءات إلى خسارة في المعلومات. لذا يتعيّن على مكاتب الإحصاءات الوطنية أن تجد التوازن بين هذه الخسائر وبين خطر الكشف. فكما تستطيع هذه المكاتب إجراء تقييم لخطر الكشف، تستطيع أيضاً تقييم خسارة المعلومات الناجمة عن اتباع مختلف تقنيات الضبط. وفيما يتعلق بالبيانات النوعية، يمكن تقييمها بوسائل مثل المقارنة المباشرة، أو مقارنة جداول الاحتمالات، أو القياس الأنتروبي. أما في حالة المتغيرات المتصلة، فيمكن قياسها عن طريق المعدلات الترتيبية، أو المعدلات المطلقة، أو تغاير المعدلات¹³. ومع ذلك، نظراً إلى أن احتمالات استخدام ملفات البيانات الجزئية واسعة للغاية، يستحيل بكل بساطة إجراء تقييم شامل لخسارة المعلومات. أما الخيار الأفضل فيتمثل في تحديد فئة المستخدمين التي تؤثر فيهم إجراءات ضبط الكشف عن الإحصاءات تأثيراً كبيراً. وعادة ما تضم هذه الفئة

13 يرجى زيارة موقع IHSN (www.ihsn.org) للحصول على مزيد من المعلومات بشأن تقنيات تقييم خسارة المعلومات.

الاطاء، 15 الاحاءات المطبقة علم، عنات السانات الحائمة المتاحة للاستخدام العام مسح عام 0002

الجزئية المتاحة للاستخدام العام (PUMA) وملفات البيانات الجزئية المتاحة للاستخدام العام المميّزة المقدّمة في القسم الثالث).

بالإضافة إلى ذلك، تعدّ عملية تشويش البيانات المتعلقة بالعمر، أو بالأحرى تعديل العمر للأشخاص المنتمين إلى عائلة، عملية ضرورية للأسر الكبيرة الحجم (10 أشخاص أو أكثر)، لأسباب حماية السرية.

وتقتطع تفاصيل المتغيرات النوعية إذا كانت الفئة المعنية لا تحترم عتبة السكان الوطنية الدنيا المحددة“.

”ستتم حماية السرية بالوسائل التالية: تبديل البيانات، إعادة تجميع القيم العليا، والعتبات الجغرافية، وتشويش البيانات المتعلقة بالسن للأسر الكبيرة الحجم، وحظر التفاصيل المقدّمة لبعض المتغيرات النوعية.

يعدّ تبديل البيانات طريقة تحدّ من الكشف عن المعلومات التي تم الحصول عليها لحماية سرية البيانات الواردة في الجداول المتكررة (عدد الأشخاص أو نسبة السكان الذين يظهرون بعض السمات). وهي تقضي بتعديل البيانات الأصلية أو تبديل تسجيلات خاصة بعينة ما. ويغال التبديل التسجيلات الفردية، كما يحمي البيانات الجزئية في الوقت عينه.

أما إعادة تجميع القيم العليا فهي إحدى طرق الحد من الكشف عن المعلومات، وهي تقضي بتجميع كل الحالات التي تتلاءم مع توزيع بتواتر معين أو تفوقها في نفس الفئة الوحيدة.

ومن ناحية أخرى، تمنع العتبات الجغرافية الكشف عن البيانات المتعلقة بأشخاص أو بأسر مرتبطة بوحدات جغرافية لا يفوق تعداد سكانها مستوى معيّناً (أنظ و صف وحدات السانات

المصدر: <http://www.census.gov/population/www/cen2000/pums/index.html>. تمت زيارة الموقع في 10 أغسطس/ آب 2010 (ترجمة من الإنجليزية)

8. هل ينبغي أن يكون النفاذ إلى البيانات الجزئية مجانياً أم مدفوع الأجر؟

إليها المستخدمون. وقد لاقى هذا البرنامج نجاحاً طفيفاً وألغى قدر كبير منه، أو اختزل إلى حد بعيد. وقد ارتفع سعر النفاذ إلى البيانات الجزئية بنسبة كبيرة، ما دفع الباحثين الكنديين إلى استخدام بيانات مجموعة في الولايات المتحدة، يسهل النفاذ إليها عن طريق اتحاد أبحاث العلوم الساسية والاجتماعية بين الجامعات (ICPSR). وقد أطلقت مبادرة إتاحة البيانات (IDD) لحل هذه المشكلة التي كانت لا تصب في المصالح الكندية. أما التحليل التالي، فكان هو شرارة التفكير في مبادرة إتاحة البيانات:

”... تتجلى الإتاحة الحقيقية في تمكين جميع المواطنين من النفاذ إلى المعلومات المعقدة ومن الحصول على المهارات اللازمة لفهم هذه المعلومات.“ وإذا كان بول برنارد Paul Brenard مدرراً للضغوطات التي تمارس على الإحصاء الكنديحتى تقلص المؤسسة تكاليفها وترفع من مداخيلها، فهو يشعر في الوقت ذاته بأنها تنعكس إلى ”نفاذ إلى البيانات في فئات المستخدمين القادرين على الدفع“. ووفقاً لبول برنارد، يرجح أن يؤدي هذا الموقف إلى ”... إعاقة مشاركة الفئات التي تملك موارد محدودة في الشأن العام، ومثلها الفئات التي تملك فرصاً ضئيلة في الاستفادة من استخدام البيانات أو تحقيق مكاسب ملموسة وفورية نسبياً منها“. وهو يضيف إن ”... من شأن ذلك أن يؤدي على المدى الطويل إلى تطور أذنى من التطور الأقصى، وإلى إتاحة غير حقيقية“ (31)

وهكذا أنشئت شبكة من 74 مؤسسة تعليمية تتيح النفاذ إلى كامل تشكيلة البيانات العامة لدى Statistique Canada. وتحتوي هذه التشكيلة على حوالي 300 ملف بيانات جزئية متاحة للاستخدام العام وغيرها آلاف ملفات البيانات، وقواعد البيانات، والملفات الجغرافية.

وتغطي رسوم الانتساب تكاليف المساعدة وتكاليف إعداد البنية التحتية التقنية حتى يستفيد منها المنتسبون والمؤسسة بأكثر قدر ممكن. إنما لا يراد منها تغطية تكاليف إنتاج البيانات.

لا تفتح بوابة النفاذ إلى مصادر مبادرة IDD إلى الوزارات التي وقع عدد كبير منها على اتفاقات مع الإحصاء الكنديتتطرق إلى موضوع المشاركة في تمويل استقصاءات معينة. وبصورة عامة، تتضمن هذه الاتفاقات على حق النفاذ إلى البيانات الجزئية. لذا لا تستجيب الخدمة المقترحة إلى حاجات الوزارات.

8.2 النفاذ مدفوع الأجر أو مجاني؟

ليس من إجابة مطلقة على هذا السؤال. فثمة حجج كثيرة تدافع عن تقليص ثمن النفاذ إلى ملفات البيانات الجزئية إلى حده الأدنى، حتى لا نقول أنها تشجع على المجانية المطلقة.

النفاذ المجاني

لعل الحجة الأساسية التي تصب في صالح مجانية النفاذ تتمثل أن الكثير من مكاتب الإحصاءات الوطنية تعزز ممارساتها في مجال نشر ملفات البيانات الجزئية. إذ تخشى هذه المؤسسات أن تبعد التعريفات المرتفعة للغاية للمستخدمين.

- يقلص النفاذ المدفوع بشكل كبير عدد المستخدمين المحتملين، وبالتالي يخفض قيمة البيانات.
- في البلدان النامية، قد يمثل عائقاً أمام المهتمين الأساسيين بالموضوع:

قد يفكر منتجو البيانات في بيع البيانات الجزئية كوسيلة لاسترداد تكاليف الإنتاج التي لا تغطيها دائماً الميزانية. لقد أصبحت مكاتب الإحصاءات الوطنية أكثر استقلالية، لذا بات يتعين عليها غالباً أن تجد أنشطة ربحية وتطبقها. وفي هذا المجال، تعدّ البيانات الجزئية منتجاً مثيراً للجدل.

إذا كان نشر ملفات البيانات الجزئية يرفع من قيمة الاستقصاء، فهو يولد أيضاً تكاليف مرتبطة بالتوثيق والإغفال. وحين لا تكون تكاليف إنتاج البيانات الجزئية المعدّة للنشر مدرجة في ميزانية الاستقصاء، لا يجد مكتب الإحصاءات الوطني طريقة أخرى لمحاولة استرجاع ما أنفقه.

عادة ما تولّد الخدمة الكاملة المتمثلة في نشر البيانات الجزئية تكاليف أخرى أيضاً متعلقة بضبط الملفات قبل النشر، ومنح تراخيص الاستخدام، واستخدام مركز بيانات محمي، والمساعدة المقدمة إلى المستخدمين، وإدارة البنية التحتية. وبصورة عامة، تستلزم هذه المهام فريق عمل جاد وبنية تحتية خاصة. وإذا كانت المؤسسة لا تملك الميزانية اللازمة أو كانت لا تستطيع تحمل مسؤولية هذه المهام اللازمة، قد يصعب تقديم خدمة جيدة وقابلة للاستمرار.

لذا يعدّ بيع البيانات طريقة تتحمّل المستخدمين من الخدمة جزءاً من تكاليفها.

8.1 مثال بلدين

لطالما كانت مكاتب الإحصاء الوطنية هي التي تتحمل تغطية التكلفة في كل أرجاء العالم. وقد عمّم ذلك عام 1980 رداً على الاقتطاعات في الميزانية وعلى ضغوطات منظمات الوصاية التي تؤدي نقل التكلفة المتعلقة بالإحصاءات من المنتج إلى المستخدم. ولا تتناول هذه الوثيقة الفحص الكامل لأنشطة مكاتب الإحصاءات الوطنية المتمحورة حول تغطية النفقات، ولكم ثمة بعض الملاحظات:

- حاول معهد الإحصاءات النيوزيلندي (SNZ) استعادة 25% من ميزانيته الإجمالية عن طريق بيع منتجات وخدمات مختلفة. وقد أخذ هذا القرار في فترة تشديد إداري في البلاد. وسمحت مبادرة المعهد بتفادي حذف بعض البرامج. غير أن المؤسسة لم تنجح في تحقيق هدفها، واعتبرت المبادرة أخيراً غير ملائمة للتنفيذ. لم يرد معهد الإحصاءات النيوزيلندي إدخال التكاليف في ميزانيته، فاعتمد سياسة لتقليص التكاليف. ووفقاً للموقع الإلكتروني الخاص بالمؤسسة، سيصبح حوالي 90% من المعلومات التي يجمعها المعهد متاحة مجاناً. وتقتطع تكاليف للتجميعات المعدة خصيصاً والجداول المفصلة. أما النفاذ إلى ملفات البيانات الجزئية فهو مدفوع الأجر، غير أن التعريفات لم تحدد بعد على الموقع الإلكتروني¹⁴.

- وضع الإحصاء الكندي برنامجاً كاملاً لتغطية تكاليف الثمانينيات لمواجهة ضغوطات الميزانية والضغوطات السياسية. وقد مضت المؤسسة في هذه المهمة على الرغم من الانتقادات الحادة التي وجهها

14 أنظر - <http://www.stats.govt.nz/about-us/making-more-informationfree/default.htm>

تمت زيارة الموقع في 18 أكتوبر/ تشرين الأول 2007 (باللغة الإنجليزية فقط)

الهامشية ضئيلة، أو حتى معدومة. ولكن الأمر نفسه لا ينطبق على المنتجات بنسختها الورقية. إذ تحدد تعرفه هذه الأخيرة وفقاً للتكاليف الهامشية لإنتاج النسخ الإضافية وإرسالها. ويمكن تطبيق المبدأ عينه على ملفات البيانات الجزئية وعلى التكلفة الزائدة التي تسببها المساعدة الإضافية المقدّمة إلى المستخدمين.

الطلاب، ومراكز الأبحاث المحلية، والجامعات، إلخ.

أما الحجة الثانية التي تصب في صالح المجانية فهي مرتبطة بصياغة حقوق النفاذ. إذ يتغير دافع فريق العمل تجاهها وفقاً لمن تتوجه (إلى مكتب الإحصاءات الوطني أو إلى مؤسسة وصاية): لذا ينبغي أن يكون النظام فعالاً ليعمل على نحو سليم.

- يستدعي النفاذ المدفوع الأجر مستوى من الجودة ومن الخدمة.
- لا يحقق النفاذ المدفوع ربحية كبيرة: إذ تصدر معظم الطلبات من الجامعات التي تملك موارد محدودة والتي تستطيع بسهولة أن تقصد مكاناً آخر وتستخدم بيانات أخرى.

وفي هذا الصدد، تظهر تجارب بعض البلدان، لا سيما من البلدان الصناعية، أن تغطية التكاليف ممكنة إلى حد ما. ولكن يبدو من الصعب مع ذلك تغطية كل التكاليف الإضافية المتعلقة بنشر البيانات الجزئية. كما يمكن إظهار أن اعتماد سياسة تعسفية لتغطية التكاليف لا تشجع على استخدام ملفات البيانات الجزئية، وتخفف من القيمة المحتملة للاستقصاء على المدى الطويل.

أما الحل المثالي لمكاتب الإحصاءات الوطنية، فيقضي بأن تدرج جميع التكاليف في ميزانية الاستقصاء، ما يعزز من قدرة النفاذ إلى البيانات إلى أقصى حد. ويمكن أن يتحمل الرعاة هذه التكاليف، ما يعزز فوائد الاستقصاء. وبعد هذا الجانب حاسماً في البلدان التي يملك فيها الباحثون إمكانيات محدودة، وحيث يملك المنتجون القليل من الموارد لإجراء تحليل البيانات.

النفاذ المدفوع الأجر

من المرجح للغاية أن يولد نظام النفاذ المدفوع الأجر مداخل. ولكن مع ذلك، ثمة عوامل أخرى ينبغي مراعاتها قبل اتخاذ القرار بوضع نظام كهذا، لا سيما الإجابات على هذه الأسئلة:

- هل يملك مكتب الإحصاءات الوطني من الناحية القانونية الحق في فرض ثمن على حقوق النفاذ إلى هذه المنتجات؟
- ما هي التكاليف التي يودّ مكتب الإحصاءات الوطني استعادتها؟
- هل يمكن تحديد هذه التكاليف بوضوح: هل سيفهمها المستخدمون ويقبلونها؟
- هل يملك المستخدمون إمكانيات الدفع؟ هل من الوارد إنشاء هيئة مستخدمين لتغطية التكاليف المتصاعدة؟ إن هذا يتطلب تحديداً واضحاً للتكاليف التي ينبغي تغطيتها، وتوزيع هذه الأخيرة بين المؤسسات المستخدمة.
- هل يمكن منح حقوق النفاذ بفعالية؟
- يمكن لأشخاص لا يتمتعون بالمهارات اللازمة لاستخدام البيانات الجزئية أن يراجعوا الملفات المعقدة المتاحة مجاناً على الموقع الإلكتروني. وقد يؤدي ذلك إلى زيادة في الحاجة إلى المساعدة.

ثمة عامل آخر مهم لا بد من أخذه في الاعتبار عند صياغة استراتيجية تعريفية: ألا وهو التناقص مع السياسة التعريفية المطبقة على منتجات وخدمات أخرى، مثل المنشورات الورقية والنفاذ إلى الإنترنت. إذ يمكن النفاذ إلى معظم المواقع الإلكترونية الخاصة بمكاتب الإحصاءات الوطنية مجاناً. حين تكون التكاليف

9. في أي وقت من دورة البث ينبغي أن تجعل البيانات الجزئية عامة؟

يتطلب إنتاج ملف بيانات جزئية بعض الوقت، ويحرك الموارد المتخصصة، كما يتطلب عملية تحقق. وعلاوة على ذلك، يتعين على مكاتب الإحصاءات الوطنية أحياناً أن تلي أهدافاً تحليلية/ علمية داخلية أو خارجية معينة. وهكذا قد تضطر إلى تأجيل إنشاء ملف بيانات جزئية بعض الأشهر بعد نشر نتائج استقصاء معين. ومهما كان جدول الأعمال المتوقع، يود الباحثون أن يتم إبلاغهم بتاريخ النشر المرتقب، حتى يتمكنوا من تخطيط دراساتهم الخاصة، ويجب أن تكون المهلة معقولة، فإذا دامت المهلة سنوات عدة، كانت النتائج أقل ثباتاً بكثير.

”حتى تتم الإتاحة بسرعة، يفضل أحياناً أن ينشر جزء من البيانات الجزئية أو البيانات المجمعة قبل تاريخ إتاحة مجموع البيانات الجزئية الكامل. (ويوصى) باستباق هذه المسائل منذ مرحلة التخطيط أو دراستها ما إن يشعر بال حاجة إليها“ [14]

يجب أن يندرج إنتاج ملفات البيانات الجزئية ونشرها في دورة البث. وينصح بشدة أن تنشر في المقام الأول المعلومات المتوجهة إلى جمهور كبير، بحيث يتمكن مكتب الإحصاءات الوطني من تلبية حاجاته الفورية ووضع نظام لمعرفة آراء الجمهور. ويشمل ذلك تقارير وصفية للاستقصاءات والتحليلات التي يجريها منتج البيانات. ومن الضروري أن يعدّ منتج البيانات الرسميون هذه الوثائق الرسمية ويبنوها/ أو ينشروها منذ بداية دورة البث.

الإطار 16 سياسة آجال نشر البيانات في المركز الوطني للإحصاءات الصحية SHCN (الولايات المتحدة)

الأخلاقية للأبحاث والمواضيع الإنسانية. وفي المبدأ، يتم إرسال البيانات هذا في هذا الإطار وفقاً لاتفاق يفرض إجراءات حماية ملائمة يتخذها الشريك المعني.

2. لا "يعترض" المركز الوطني للإحصاءات الصحية سبيل البيانات الجاهزة للإتاحة. كما لا يقدم إلى شركائه حق النفاذ المميز والمسبق إلى ملفات البيانات أو الجداول الجاهزة للإتاحة. وأخيراً، لا يقدم حق النفاذ المميز إلى الجداول المستندة إلى ملفات بيانات لم تصبح متاحة للعام بعد. عندما ترسل بيانات غير منشورة إلى مستخدم، يستحسن دراسة إتاحتها إلى مقدمي طلبات آخرين، في ما لا يتجاوز بنود السرية. أما ملفات الجداول التي لم تتم الموافقة على نشرها بعد نظراً إلى خطر الكشف عن المعلومات السرية التي تحتوي عليها، فيمكن مراجعتها لاحقاً في مركز البيانات التابع للمركز الوطني لإحصاءات الصحة (أو بما يتوافق مع سياسة المركز الوطني للإحصاءات الصحية في مسائل السرية، وذلك في عقود استخدام مميزة)، من أجل ضمان أوسع نطاق ممكن لاستخدام البيانات المنتجة. وفي حالات نادرة (كحالات منشورات وزارية يعود تاريخ إصدارها إلى زمن بعيد مثلاً)، يمكن أن تتاح البيانات المجدولة قبل إتاحتها للعام.

أما استثناءات هذه القاعدة فهي عادة نادرة وتبرر حالة بحالة. ويجب أن تقدم جميع الطلبات قبل المباشرة في عملية جمع البيانات إلى المسؤول عن مسائل السرية التي يحددها المركز الوطني للإحصاءات الصحية، وأن يوافق يوافق عليها المدير العام". (ترجمة من الإنجليزية)

هذا الصدد، تقضي سياسة المركز الوطني للإحصاءات الصحية بنشر البيانات الجزئية في أبكر وقت ممكن بعد جمع البيانات، وتخضع فقط للقيود التي تفرضها عليها مواردها المتاحة، والقيود التقنية والمتطلبات النوعية. ولا يعيق المركز الوطني للإحصاءات الصحية الإتاحة السريعة للبيانات الجزئية ليحصيها فقط في مساعدته، أو شركائه، أو فريق عمل مؤسسات أخرى تملك حقوق النفاذ إليها.

1. يتم وضع ملفات البيانات الجزئية المتاحة للاستخدام العام في التصرف بعد إنهاء المراحل الضرورية للتحرير، والضبط، والموافقة لدى السلطات المختصة، لا سيما لجنة ضبط السرية التابعة للمركز الوطني لإحصاءات الصحة. وفقاً للمصلحة التي تمثلها البيانات ومجال الاختصاص، قد يسخر المركز الوطني للإحصاءات الصحية بعض الشركاء (منهم الممولين) في إطار عملية إعداد البيانات التي تسبق إتاحتها، بما في ذلك تعديل الملف، وإعادة ترميزه، وتحديد هيكلته النهائية. وإذا تبين أن هذه المراحل مهمة، قد يدرس المركز الوطني للإحصاءات الصحية إمكانية إرسال بيانات لم تنشر بعد للعام إلى شريك له. أما إرسال البيانات غير الجاهزة بعد الإتاحة في هذا الإطار، فينبغي أن يتم استباقه في أحكام سياسة السرية التي يتبعها المركز. كما ينبغي أن يتم الإرسال بما يحترم الأحكام القانونية التي يطبقها المركز، ومع الموافقة الخطية بعد العلم بالمخاطر والتوصيات الصادرة عن اللجنة

المصدر: http://www.cdc.gov/nchs/about/policy/data_release.htm (تمت زيارة الموقع في 7 يونيو/ حزيران 2010، بالإنجليزية فقط)

10. ما هي المتطلبات اللازمة من حيث البنية التحتية التقنية؟

- تقييم مدى ملاءمة المعلومات المستقاة مع حاجات الباحث المعني (مثل عالم الاستقصاء والمفاهيم والتعريفات). وتضطلع البيانات الوصفية بهذا الدور إذ تشكل بيانات الملف التوثيقية.

الإطار 17 مجموعة أدوات إدارة البيانات الجزئية (الشبكة الدولية لاستقصاءات الأسر المعيشية IHSN)

طوّرت خدمات البيانات الترويجية للعلوم الاجتماعية NSD وفريق البنك الدولي لمعالجة البيانات التنموية في إطار الشبكة الدولية لاستقصاءات الأسر المعيشية IHSN مجموعة أدوات إدارة البيانات الجزئية Management Toolkit Microdata (المشار إليها أدناه باسم مجموعة الأدوات). وتهدف مجموعة الأدوات إلى الترويج لاعتماد المعايير الدولية والممارسات الفضلى عند توثيق البيانات الجزئية وبثّها وحفظها.

وتشتمل مجموعة الأدوات على نموذجين. إذ يُستخدم محرر البيانات الوصفية Metadata Editor لتوثيق البيانات وفقاً للمعايير الدولية النافذة الخاصة بالبيانات الوصفية. (مبادرة إتاحة البيانات DDI ومعياري دبلن كور Dublin Core). ويمكن للمستخدمين بفضل برنامج المستكشف Explorer المجاني أن يقرأوا الملفات التي يستحدثها محرر البيانات الوصفية. كما يتيح هذا البرنامج للمستخدمين عرض البيانات الوصفية وإعادة تصدير البيانات في صيغ مشتركة مختلفة (مثل Stata وSPSS وغيرها). ويعوّل كل من محرر البيانات الوصفية والمستكشف على تكنولوجيا نيسستار Nesstar وقد أنشأتها خدمات البيانات الترويجية للعلوم الاجتماعية NSD. وأخيراً يتم استخدام منشئ القرص المضغوط CD-ROM Builder و NADA Portal لتطوير منتجات سهلة الاستخدام (قرص مضغوط وموقع إلكتروني) بغية نشر البيانات وأرشفتها.

انظر <http://www.ihsn.org/toolkit>

- النفاذ إلى البيانات. تتطلب هذه الوظيفة نظام استخراج للبيانات و/أو تحميلها. ويمكن عامة تحميل الملفات المعنية عبر موقع إلكتروني أو مدخل الموقع أو خادم FTP. ويستطيع مكتب الإحصاءات الوطني استخدام نوع الأداة هذا داخلياً بغية تقديم البيانات على قرص مضغوط / قرص فيديو رقمي.

- استخدام البيانات. إذ لا تتوافر أداة واحدة لتنفيذ أعمال الباحثين التحليلية، ويؤثر هؤلاء الاستناد إلى صيغ متنوّعة، ما يتيح لهم الاستعانة بالأدوات التي يقع عليها خيارهم. ومن الأدوات الشائعة، نذكر SPSS وSTATA وSAS وASCII.

ندرج أدناه خصائص نظام كفوء لفهرسة البيانات الجزئية:

- من وجهة نظر المستخدم، يكون الفهرس جيداً عندما:
 - يتوافق مع معيار بيانات وصفية دولية.
 - إذ تسهّل بشكل ملحوظ معايير البيانات الوصفية الدولية في صيغة XML إنتاج الفهارس وصيانتها.
 - يستند إلى الموقع الإلكتروني لتسهيل الأبحاث.

”يتطلب النفاذ إلى البيانات الجزئية والاستفادة القصوى من هذه البيانات بنية تحتية تكنولوجية متكيفة وتقارباً واسعاً في وجهات النظر على المستوى الدولي في كل من العمل البيئي وآليات مراقبة نوعية البيانات الفعّالة. (...) وتكتسي استدامة البنية التحتية المطلوبة للنفاذ إلى البيانات أهمية خاصة. لذا يتعين على مؤسسات الأبحاث والوكالات الحكومية أن تطلع رسمياً بمسؤولية التأكد من أن البيانات يتم حفظها وإدارتها بفعالية بالإضافة إلى جعلها متاحة للعامة بغية استخدامها بشكل فعال وملائم على المدى الطويل. ولا بدّ [علاوة على ذلك] من تشجيع استخدام التقنيات والأدوات الرامية إلى ضمان سلامة البيانات البحثية وأمنها. ومن أجل ضمان سلامة البيانات البحثية، يجب تسخير الجهود كافة من أجل التأكد من اكتمال البيانات وعدم اشتغالها على أخطاء، أما بالنسبة لأمن البيانات والبيانات الوصفية والتوصيفات الملائمة، فمن المهم حماية البيانات من فقدان التلف والتعديل والنفاذ غير المصرّح به سواء كانت كلها مقصودة أو غير مقصودة، وذلك بما يتوافق مع بروتوكولات أمن البيانات الصّريحة“ [17]

ولا بدّ من وضع بنية تحتية متكيفة بغية تغطية كل الأوجه المختلفة الخاصة بأرشفة البيانات الجزئية (من توثيق البيانات وفهرستها وبثّها وإغفالها وحفظها).

توثيق البيانات الجزئية

تمّ وضع معايير للبيانات الجزئية الدولية بغية إضفاء طابع رسمي على توثيق البيانات الجزئية والموارد ذات الصلة. وتشكل مبادرة إتاحة البيانات DDI ومعياري دبلن كور Dublin Core المفصلتان في الفصل الثاني للحلول العملية. فيكون من الأسهل تحقيق توثيق البيانات بما يتلائم مع هذه المعايير وذلك بفضل محرري البيانات الوصفية المختصين وهما مجموعة أدوات إدارة البيانات الجزئية من الشبكة الدولية لاستقصاءات الأسر المعيشية IHSN (الإطار 17) وبرمجية إدارة البيانات Nesstar Publisher الخاص بخدمات البيانات الترويجية للعلوم الاجتماعية NSD.

فهرسة البيانات الجزئية وبثّها

لا بدّ من إعلام المستخدمين المحتملين بطريقة صحيحة عن توافر خصائص لمجموعات البيانات. إذ يملك عدد منهم قليلاً من المعلومات حول مجموعات البيانات المتاحة — ولا يعدّ ذلك كافياً، فمن الضروري تزويد المستخدمين ببيانات وصفية نوعية، ومن المفصّل أن تكون على شكل فهرس للبحث على الشبكة.

يهدف فهرس البيانات الجزئية إلى السّماح للمستخدمين بالنفاذ بسهولة إلى البيانات والتوثيق وذلك عبر صيغة عملية جدّاً. ويحتوي فهرس الاستقصاء على أدوات متنوّعة من أجل تأدية الوظائف التالية:

- إيجاد ملف البيانات الذي يلبي أكثر من غيره احتياجات المستخدم. ولا يعتبر نوع الأداة هذا مهماً جداً عندما تكون كمية ملفات البيانات الجزئية محدودة، إلا أنه يصبح غاية في الأهمية عندما يتزايد عدد الملفات. ففي هذه الحالة، يوفّر أداة تتيح البحث عن ملفات البيانات وفقاً للمتغيّرات.

- يكون غنياً بالبيانات الوصفية بما فيها المتغيرات.
 - إذ تصبح فهارس الاستقصاء أكثر استدامة وفعالية عندما تحتوي البيانات الوصفية على توصيف مفضل للاستقصاء نفسه (معلومات حول العنوان، والباحث الرئيس، والعينة، وتاريخ جمع البيانات، والمواضيع المدروسة، والتغطية الجغرافية، وغيرها) وإنما أيضاً على مستوى كل متغير (تسمية المتغيرات وبطاقة التعرّف، والفئات، وصياغة الأسئلة، والتوجيهات للباحثين، والتعريفات). يمكن الاطلاع على معيار مبادرة إتاحة البيانات DDI وأدوات البرمجيات التي تقدّمها الشبكة الدولية لاستقصاء الأسر المعيشية IHSN ولاسيما مجموعة أدوات البيانات الجزئية ومنصة NADA (متوفرة على موقع www.ihsn.org).
 - يكون متاحاً للبحث في حقول الدراسة كافة. في إطار معيار مبادرة إتاحة البيانات DDI، يفرض ذلك للقول إنه يتعين على الفهرس أن يتيح القيام بأبحاث حول دراسة ما (العنوان، والسنة، والبلد، والمنظمة) وحول المتغيرات (تسمية المتغيرات وبطاقة التعرّف وبطاقة تعرف قيمة المتغير). ومن المستحسن أن يشتمل الفهرس على وظائف للبحث في مجمل النص على أن تكون سهلة الاستخدام.
 - يتضمن معلومات واضحة حول سياسة النفاذ إلى البيانات والإجراءات المتبعة.
 - يتضمن قائمة بالوثائق المرجعية والروابط التي تؤدي مباشرة إليها (الاستمارات والمنهجية والكتيبات والتقارير)
 - يتيح إمكانية "البحث بالكلمة المفتاح" وفقاً لتصنيف معياري.
 - وقد وضع مجتمع أرسفة البيانات مكنزاً يصف الموضوعات التي تغطيها مجموعات البيانات المدرجة في الفهارس الخاصة بها. والمكنز عبارة عن قائمة من المصطلحات والمفاهيم المستخدمة لوصف عناصر محددة (مجموعات البيانات، والمتغيرات، والكتب، وغيرها). وعادةً ما يتم تنظيم مصطلحات المكنز على غط الشجرة أو على شكل قائمة متسلسلة هرمياً (بدءاً بالمصطلحات العامة للانتقال إلى مفاهيم محدّدة أكثر). وبشكل عام، يظهر المكنز المرادفات والمصطلحات ذات الصلة فيضمن أن تحقق أبحاث المستخدمين نتيجة حتى وإن لم يستخدم هؤلاء المصطلحات المخصّصة.
 - ويستند عدد كبير من الأرشيف على مكنز لإضافة الكلمات المفتاح المرتبطة بدراسة أو مفاهيم خاصة بالمتغيرات. ويعزّز استخدام المكنز الاتساق إذ يضمن استخدام المصطلح عينه لوصف غرض معيّن. فضلاً عن ذلك، حين يصبح المكنز متاحاً أمام المستخدمين للبحث عن البيانات، قد تزيد فرص استخدامهم للمصطلحات والمفاهيم التي تساعدهم على الحصول على قائمة النتائج الأكثر دقةً.
 - على سبيل المثال، يعتمد الفهرس الذي تديره المنظمة الأوروبية لأرسفة البيانات في العلوم الاجتماعية CESSDA، على مكنز.
 - أتاح الفهرس عرض نتائج الأبحاث بسرعة حتى لو كان ضخماً، ما حتمّ إذا الاستعانة بنظام فهرسة فعّال.
 - قدّم الفهرس أداة للمقارنة بين منتجات الفهرس. وتعتبر هذه الوظيفة مفيدة بغية مقارنة المتغيرات بين استقصاءات معيارية متنوّعة أو للمقارنة عند تنزيل نسخات مختلفة لاستقصاء واحد.
 - قدّم الفهرس معلومات واضحة عن سياسات النفاذ المطبقة في كل دراسة (مثل توافر البيانات الجزئية، وعند الاقتضاء توافر توصيف واضح لأمط الحصول عليها).
 - قدّم الفهرس وظائف جيّدة وفعّالة للمساعدة على الشبكة.
 - قدّم روابط بين منتجات الفهرس والموارد المتاحة على موقع إلكترونيّ خارجيّ وإذا سمح بإضافة معلومات مكّملة مثل المراجع البيبليوغرافية إلى المنشورات القائمة على دراسة ما.
 - ومن وجهة نظر مدير الفهرس، يكون نظام الفهرسة جيّداً إذا:
 - قدّم بيئة آمنة لتخزين البيانات والبيانات الوصفية ومشاركتها.
 - وفّر أدوات إدارة عملية النفاذ إلى البيانات الجزئية (من الموافقة التلقائية للبيانات الجزئية التي يمكن النفاذ إليها بدون قيود إلى أنظمة إدارة إجراءات النفاذ المشروطة بموافقة ومعالجتها).
 - وجد النظام حلاً لإيصال ملفات متاحة للاستخدام العام وأخرى خاضعة لترخيص.
 - شكّل النظام وسيلة مضمونة لإيصال البيانات الجزئية والوثائق، ما يزيد معدل نفاذ المستخدمين النهائيين.
 - جمع النظام معلومات حول مستخدمي الفهرس والتنزيلات التي يقومون بها وفي حال اقتضى الأمر جمع أيضاً معلومات حول غايات استخدام البيانات. ويثير تسجيل مشابه اهتمام رعاة الدراسات إذ يسمح لهم بتقييم استخدام البيانات الجزئية. كما وأنّه يعتبر مهماً للمستخدمين النهائيين إذ يصبح بالإمكان إبلاغهم عند نشر نسخات جديدة من البيانات أو عند مراجعة الدراسات التي نزلوها.
- ### إغفال البيانات الجزئية
- يستدعي إغفال البيانات تواجد فريق عمل كفوء في مجال الإحصاءات على أن يألف استخدام برمجيات Stata أو SPSS. وتتيح بعض البرمجيات المختصة قياس مخاطر الإفصاح عن البيانات أو التخفيف منها، ولكن، لا تقدّم أي تطبيقات حلولاً متكاملة ومرضية للملفات ذات البيانات المعقدة المتسلسلة هرمياً. ومن الناحية العملية، يبقى إغفال البيانات إلى حدّ كبير، عملية تتلائم مع الحالة المطروحة. وحالياً، يتمّ العمل على تطوير أدوات وخطوط توجيهية كفيّلة بتسهيل إغفال البيانات (من بين أعمال أخرى في الشبكة الدولية لاستقصاء الأسر المعيشية IHSN).

وكما رأينا، يتمّ إغفال ملف الاستقصاء على مرحلتين: إذ لا بدّ أولاً من تحديد عناصر من المحتمل أن تكون عرضة لخطر الإفصاح، ومن ثمّ وضع تقنيات تقييد للبيانات أو تشويشها بغية التقليل من هذا الخطر. وتستلزم هذه المرحلة الأخيرة تدخّل شخص يتمتع بالكفاءات اللازمة حتى يتمكن من فرض قيود على المعلومات التي تضرّ بحد أدنى بمستخدمي الملفات النهائيين.

حفظ البيانات الجزئية والبيانات الوصفية

تعتبر البيانات والبيانات الوصفية الرقمية عرضة لتقادم البرمجيات ومكوّنات المواد ووسائط التخزين كما وأنها عرضة للتهديدات المادية والأخطاء البشرية. لذا يتطلب حفظ البيانات والبيانات الوصفية على المدى الطويل، وضع إجراءات وبنية تحتية ملائمة. وتفصّل ورقة عمل للشبكة الدولية لاستقصاءات الأسر المعيشية أعدّها الاتحاد بين الجامعات للأبحاث السياسية والاجتماعية المبادئ والممارسات الفضلى في مجال حفظ البيانات. [8]

11. ما هي المتطلبات المؤسسية المتعلقة بإتاحة ملفات البيانات الجزئية؟

المهنية

يجب أن تستند التدابير المؤسسية لإدارة بيانات البحث على معايير مهنية قابلة للتطبيق وعلى قيم مدرجة في قواعد السلوك الحسن في الأوساط العلمية المعنية. (...)

التوافقية

تعدّ التوافقية التكنولوجية والمعنوية ضرورية لتسهيل وتشجيع النفاذ إلى بيانات البحث واستخدامها في سياق دولي متداخل التخصصات. ويفترض بتجارب النفاذ أن تأخذ في الاعتبار على نحو رسمي المعايير الدولية الثابتة والمطبقة في ما يتعلق بتوثيق البيانات (...).

النوعية

ترتبط قيمة البيانات وفائدتها إلى حد كبير بنوعية هذه البيانات. لذا يتعين على مديري البيانات ومؤسسات جمعها أن يحرصوا حرصاً شديداً على احترام معايير النوعية الواضحة. (...)

الأمن

ينبغي الحرص بالأخص على تشجيع استخدام التقنيات والأدوات الرامية إلى ضمان سلامة البيانات وأمنها (...).

الكفاءة

يتمثل أحد الأهداف الأساسية الملائمة لتشجيع النفاذ إلى البيانات ومشاركتها في تحسين الكفاءة العالمية لعملية (جمع البيانات) التي تمولها الأموال العامة، وذلك في سبيل تفادي التكرار غير المجدي والمكلف لجهود جمع البيانات. (...)

واجب المساءلة

يجب أن يخضع عمل تدابير النفاذ إلى البيانات إلى تقييم دوري تجريه فرق المستخدمين، والمؤسسات المسؤولة، ومؤسسات التمويل (...). (17)

ويهمّ احترام هذه المبادئ من دون شك بتدابير جديدة وذهنيات جديدة. إذ تعرّف منظمة التعاون والتنمية الاقتصادية OECD، في خطوطها الإرشادية (17)، الأسئلة الكبرى المتعلقة بمنح حقوق النفاذ إلى البيانات (وتنطبق هذه الأسئلة أيضاً على البيانات الجزئية التي ينشئها منتجو البيانات الرسمية لغايات إحصائية):

• "الأسئلة المؤسسية وأسئلة الإدارة: على الرغم من إمكانية نفاذ أفضل إلى البيانات تشكل مسألة هامة بالنسبة إلى كل الجماعات العلمية، فإن تنوع الأنشطة العلمية يحيلنا إلى الاعتقاد بأن المقاربات المصممة خصيصاً والنماذج المؤسسية المتنوعة لإدارة البيانات ستكون فعالة لا شك للاستجابة لحاجات الباحثين.

• الأسئلة المالية والأسئلة المتعلقة بالميزانية: تستلزم البنية التحتية الخاصة بالبيانات العلمية تخطيطاً ميزانياً خاصاً متواصل ودعماً

تعدّ عملية إتاحة ملفات البيانات الجزئية بالنسبة إلى كثير من مكاتب الإحصاءات الوطنية نشاطاً حديثاً للغاية، من شأنه استحداث وسائل استخدام هائلة وهامة لبياناتها. وفي وثيقة بعنوان المبادئ والخطوط الإرشادية لمنظمة التعاون والتنمية الاقتصادية OECD بشأن النفاذ إلى بيانات البحث التي تمولها الأموال العامة، حددت منظمة التعاون والتنمية الاقتصادية OECD سلسلة من المبادئ الرئيسية الموجهة إلى مقدّمي البيانات:

"الانفتاح

يعني بالانفتاح حق النفاذ إلى البيانات على وجه المساواة في قلب جماعة الباحثين الدولية، بأدنى تكلفة ممكنة، ويفضّل ألا تتعدى هذه الأخيرة التكلفة الهامشية للإتاحة. (...)

المرونة

تفترض المرونة أن تؤخذ في الاعتبار تطورات سريعة في عالم تقنيات المعلوماتية عادة ما لا يمكن التنبؤ بها، (...). إلى جانب الأنظمة القانونية والثقافية الخاصة بكل بلد (...).

الشفافية

يجب أن تكون المعلومات المتعلقة ببيانات البحث والمؤسسات المنتجة للبيانات، إلى جانب توثيق البيانات وميزات الشروط التي تحكم استخدامها، متاحة على الصعيد الدولي، بكل شفافية، ويفضّل أن يكون ذلك عبر شبكة الإنترنت. (...)

الامتثال للقوانين

يجب أن تحترم أحكام النفاذ إلى البيانات الحقوق والواجبات الشرعية الخاصة بكل أطراف نشاط البحث العام (...).

حماية الملكية الفكرية

يجب أن تراعي شروط النفاذ إلى البيانات ظروف تطبيق حقوق الملكية أو غيرها من التشريعات المتعلقة بالملكية الفكرية والتي يمكن أن تطال قواعد بيانات البحث التي تمولها الأموال العامة. (...)

المسؤولية الرسمية

ينبغي أن تشجع تدابير النفاذ الممارسات المؤسسية الواضحة والرسمية، مثل وضع القواعد والأنظمة المتعلقة بمسؤوليات مختلف الأطراف المشاركة في أنشطة البيانات. وينبغي أن تعنى هذه الممارسات بحقوق تأليف البيانات، وذكر اسم منتجها، والملكية، والإتاحة، وقيود الاستخدام، والتدابير المالية، والقواعد الأخلاقية، وشروط الترخيص، والمسؤولية المدنية، والأرشفة المستدامة. (...)

مالياً ملائماً. ولا يمكن تعزيز استغلال بيانات البحث بأفضل الطرق إذا أضيفت تكاليف النفاذ والإدارة والحفظ إلى مشاريع البحث. وينبغي التشديد على أن تكاليف تخزين البيانات وإدارتها انخفضت كثيراً في السنوات الأخيرة، وأن نقص المعلومات بشأن هذه التغييرات بحد ذاته يمكن أن يشكّل عائقاً أمام التقدم.

- الأسئلة القانونية والسياسية: تؤثر التشريعات الوطنية والاتفاقات الدولية، لا سيما في مجالات قوانين الملكية الفكرية وحماية الخصوصية على سبيل المثال، مباشرة في الممارسات المتعلقة بالنفاذ إلى البيانات ومشاركتها، وينبغي أخذها في الاعتبار بأكملها في صياغة أحكام النفاذ إلى البيانات.
- أسئلة الثقافة والسلوك: يمثل وضع آليات تدريب ومكافئة ملائمة جزءاً مهماً من عملية التشجيع على الممارسات المتعلقة بالنفاذ إلى البيانات ومشاركتها، وتنطبق هذه الاعتبارات على من يمول البيانات، ومن ينتجها، ومن يديرها ويستخدمها (...).

ويجب أن تذكر مسؤولية مختلف جوانب النفاذ إلى البيانات وإدارتها في وثائق تعنى بالمهام الرسمية التي تضطلع الخاصة بالمعاهد، وطلبات المساعدات، و عقود الأبحاث، واتفاقات النشر والتراخيص على سبيل المثال (17).

المرجعيات الرقمية الموثوق بها تشكل بديلاً ممكناً

بالنسبة إلى بعض منتجي البيانات، يعدّ وضع وإدارة أرشيف بيانات وخدمة إتاحة لها هدفاً بعيداً عن الواقعية، وذلك لأسباب عدة من بينها أسباب ميزانية وقانونية. أما إحدى البدائل فتتمثل في تفويض هذه المهمة. ولأخذ مثال أرشيف بيانات المملكة المتحدة¹⁵ UKDA الذي يتخذ مقره في قلب جامعة إيسيكس Essex، والذي يدير ويتيح بيانات معاهد الإحصاءات، ومؤسسات البحث، والباحثين أنفسهم، وثمة مثال آخر هو: اتحاد أبحاث العلوم الساسية والاجتماعية بين الجامعات (ICPSR) القائم في جامعة ميتشيغان والذي يشغل وظيفة مماثلة في الولايات المتحدة¹⁶.

لا تحقق أرشيفات البيانات إدارة فعالة لعملية منح التراخيص فحسب، بل تؤدي أيضاً دوراً محورياً في رفع قيمة البيانات والابتكار. فعلى سبيل المثال، لنذكر الصفحات الإلكترونية الجديدة الخاصة بأرشيف بيانات المملكة المتحدة، التي تقدّم التوجيه فيما يتعلق بإدارة البيانات ومشاركتها. أما الهدف فيتمثل في أن تعرض على مؤلفي البيانات ومديريها والقيمين عليها الاستراتيجيات والطرق الفضلى الرامية إلى إنشاء مجموعات بيانات معدّة للاستفادة العامة¹⁷.

15 أنظر <http://www.data.archive.ac.uk/>

16 أنظر <http://www.esds.ac.uk/aandp/create/research.asp> أو <http://www.nsf.gov/pubs/2005/nsb0540/>

17 أنظر <http://www.data.archive.ac.uk/sharing/>

12. ما السبيل إلى التشجيع على استخدام ملفات البيانات الجزئية؟

والتعليم¹⁹. وفي كندا، سبق وأن تعرّف الكثير من الباحثين إلى ملفات البيانات الجزئية. غير أن الأمر ذاته لم يسر على المؤسسات المكلفة بتحقيق الترويج. وقد تحسّن الوضع بفضل نظام تراسل إلكتروني يسمح للباحثين، وفريق عمل مكاتب الإحصاءات الوطنية، والوسطاء، وغيرهم من الأشخاص المهتمين بطرح الأسئلة وبمشاركة خبراتهم. كما يشكّل أرشيف الإجابات المقدمّة مصدراً نفيماً للمعلومات ويخفف الضغط عن مكاتب الإحصاءات الوطنية التي لا تضطر إلى الإجابة على جميع الأسئلة التي تتلقاها.

وأخيراً، لعلّ أحد مفاتيح نجاح بث ملفات البيانات إلى الباحثين هو تفاعل فريق عمل مكاتب الإحصاءات الوطنية مع شبكة جماعة الباحثين والأرشيفات. ولا شك أن ذلك يساعد الطرفين على التحديد المتبادل للحاجات والمعضلات، وينشئ قاعدة تواصل في الاتجاهين

هل يعدّ وجود ملفات البيانات الجزئية المتاحة للاستخدام العام أو الملفات الخاضعة لترخيص كافياً لتوسيع قاعدة مستخدمي هذه المنتجات؟ للأسف، ليس بالضرورة. إذ يتطلّب ذلك إقناع المستخدمين بالمشاركة، أي إعدادهم لهذه الغاية.

تمثّل الاستقصاءات الوطنية عن طريق استطلاعات الرأي، وملفات البيانات المتاحة للاستخدام العام والناجمة عن المسوح مواضيع تثير اهتمام جماعة واسعة من الباحثين والمحليلين. ويفترض بالتوثيق والدعاية الملائمة أن يساهما في استخدام مجموع هذه البيانات على نطاق واسع للغاية. وفي هذا الصدد، يظهر البرنامج العالمي للاستقصاءات الديموغرافية واستقصاءات الصحة (Health Survey (DHS, Demographic and) بصورة واضحة ضخامة حجم الطلب على مجموع البيانات من هذا النوع¹⁸. أما مجموع بيانات برنامج DHS، فهي متاحة مجاناً وبسهولة. وقد ساهمت مجموعات البيانات هذه، بعد أن حملها عدد كبير من المستخدمين، في إصدار سلسلة غنية ومتنوعة للغاية من الدراسات والمنشورات.

في البلدان التي اعتادت إنتاج هذا النوع من الملفات، أصبح تعدد الآراء في مناقشة السياسة استناداً إلى هدف استخدامها أمراً طبيعياً. غير أن استخدام ملفات البيانات الجزئية قد لا يكون مفهوماً على نحو صحيح في البلدان الأخرى. وقد تضطر الجهات التي بدأت للتو في بث ملفات البيانات الجزئية إلى اعتماد طرق متنوعة للتشجيع على استخدامها وإعلام المستخدمين بقيمتها، وبتحديد أهدافها أيضاً.

ولا بدّ أن تؤدي ثقافة مشاركة البيانات والتعاون في مجالها في إغناء معارف جديدة. لذا يطلب من مكاتب الإحصاءات الوطنية وشركائها أن تروّج لاستخدام ملفات البيانات الجزئية، وذلك عن طريق الندوات والدورات التدريبية مثلاً، على الصعيد الوطني والدولي. فكثيرة هي الفرص لتحقيق ذلك.

تعدّ ملفات البيانات الجزئية مهمة في مجالي البحث والتعليم. كما يعدّ استخدامهما لأغراض إعداد السياسات والبرامج مهماً بالنسبة إلى الحكومات الوطنية والوكالات الدولية على السواء. ويفترض أن تعتبر هذه الأخيرة حليفة لمكاتب الإحصاءات الوطنية في دعم الاستخدام السليم لملفات البيانات الجزئية والترويج له. كما تستطيع الجامعات أن تؤدي دوراً أساسياً في تدريب مستخدمي جدد.

من الضروري أن يطلع الأشخاص المؤهلون على وجود المنتج وأن يحيطوا علماً بالفوائد التي يقدّمها، كما يجدر تنظيم حملات توعية لضمان نجاح بث بيانات مكاتب الإحصاءات الوطنية. ويجري ذلك عن طريق تحديد المؤسسات والأشخاص المنتسبين إليها الذين قد يصبحون مستخدمين محتملين، والتقرّب منهم. وفي حالات كثيرة، تدرك مكاتب الإحصاءات الوطنية هذه الضرورة، وتتلقى طلبات من هؤلاء المستخدمين. وعلاوة على ذلك، تشكّل الكتيبات والدعوات إلى الندوات بعض الإجراءات التي قد يتم اللجوء إليها.

وكما ذكرنا سابقاً، ينبغي أيضاً تدريب مستخدمين محتملين. إذ تنظم مبادرة إتاحة البيانات الكندية في هذا الإطار دورات تدريب هدفها الترويج لاستخدام ملفات البيانات الجزئية وغيرها من المنتجات ودعمها، لغايات البحث

19 ندعو القراء إلى زيارة موقع تقديم وثائق التدريب الخاص بمبادرة IDD: <https://ospace.scholarsportal.info/handle/1873/69> (تمت زيارة الموقع في 18 أكتوبر/ تشرين الأول 2007)

الملحق 1

طلب نفاذ إلى مجموع البيانات الخاضعة لترخيص في إطار مشروع بحث محدد

يجب تكييف نموذج الطلب هذا مع المتطلبات الخاصة المطبقة.

ستبقى المعلومات التي تقدّمونها في نموذج الطلب هذا سرية، إلا في حال انتهاك الاتفاق القانوني المبرم، إذ يبلغ مكتب الاستقصاءات الوطني في هذه الحالة معاهد الإحصاءات الشريكة القائمة في بلدان أخرى بالأمر. الرجاء إرسال نموذج الطلب هذا عن طريق البريد أو الفاكس إلى العنوان أو الرقم التالي، مرفقاً برسالة من المؤسسة المسؤولة:

العنوان: [العنوان]

الفاكس: [رقم الفاكس]

البريد الإلكتروني (نسخة مصورة): [عنوان البريد الإلكتروني]

العنوان والرقم الخاصين بمجموع (أو مجموعات) البيانات المطلوبة (العنوان الكامل، والسنة، ورقم المرجع كما في فهرس الاستقصاء):

الشروط المطبقة

في هذا الاتفاق،

1. يمثل "الباحث الرئيسي" الباحث المكلف الذي سيتم التواصل معه بشكل رئيسي في ما يخص هذا الاتفاق. ويتحمل الباحث الرئيسي مسؤولية احترام شروط هذا الاتفاق في ما يتعلق بالنفاذ إلى البيانات. وينبغي أن يكون الباحث الرئيسي مؤهلاً لتمثيل الجهة المستفيدة في إطار هذا النوع من الاتفاقات.

2. يمثل "الباحثون الآخرون" الأشخاص غير الباحث الرئيسي، ومن بينهم المساعدون، الذين سيتمكنون من النفاذ إلى البيانات السرية.

3. تمثل "الجهة المستفيدة" المنظمة/الجامعة/المؤسسة التي توظف الباحث الرئيسي.

القسم أ. الباحث الرئيسي

- الاسم
- الكنية
- المنصب
- المنظمة
- الوظيفة في المنظمة
- العنوان البريدي
- الهاتف (مع رمز البلد)
- الفاكس (مع رقم البلد)
- البريد الإلكتروني

القسم ب. الباحثون الآخرون

الرجاء ذكر أسماء كل الأعضاء الآخرين في فريق الباحثين المخوّل النفاذ إلى البيانات السرية، ومناصبهم والمنظمات التي ينتمون إليها.

- الإسم والكنية
- الوظيفة
- المنظمة الحاضنة

القسم ج. الجهة المستفيدة

اسم المنظمة

نوع المنظمة (اختر الإجابة المناسبة)

- وزارة / إدارة عامة
- جامعة
- مركز أبحاث
- مؤسسة خاصة
- منظمة دولية
- منظمة غير حكومية (وطنية)
- منظمة غير حكومية (دولية)
- غير ذلك (يرجى التحديد)

موقع المنظمة الإلكتروني (رابط URL)

العنوان البريدي

القسم د. التعريف بالاستخدام المنتظر للبيانات

يرجى وصف مشروع البحث (الأسئلة، الأهداف، الوسائل، النتائج المتوقعة، الشركاء). وإذا كانت المعلومات المقدّمة غير كافية، يمكن أن يرفض طلبكم أو أن تلتقوا طلباً بمعلومات مكمّلة. ويمكن أن ترفق هذه المعلومات في ملحق لهذا الطلب.

قائمة النتيجة المنتظرة (النتائج المنتظرة) وسياسة البث

القسم هـ. تحديد البيانات والمتغيرات اللازمة

تقدّم مكاتب الإحصاءات الوطنية على مواقعها الإلكترونية بيانات وصفية مفصلة، وهي تضم وصفاً لملفات البيانات والمتغيرات التي تكوّن كل مجموعة من البيانات. أما الباحثون الذين لا يحتاجون إلى الاطلاع على مجموع البيانات بكامله، فيرجى منهم تحديد المجموعة الفرعية للمتغيرات أو الحالات التي تهمهم. وهكذا يقلّص خطر الكشف عن البيانات، وبالتالي تزيد فرص الحصول على البيانات المطلوبة.

يرمي الطلب إلى النفاذ إلى (اختر الإجابة الملائمة):

- مجموع البيانات الكاملة (كل الملفات والحالات)
- مجموعة فرعية من المتغيرات و/أو الحالات، بما يتوافق مع ما يلي (ملاحظة: تحتوي المجموعات الفرعية بشكل منهجي على المتغيرات مثل معاملات التزجيج الخاصة بالعينة، ومعرّفات التسجيلات)

القسم و. اتفاق النفاذ إلى البيانات

يوافق الباحث الرئيسي والباحثون الآخرون على الامتثال إلى الشروط التالية:

1. يحصر الاطلاع على البيانات السرية بالباحث الرئيسي والباحثين الآخرين المحددین في هذا الاتفاق.

2. يمنع القيام بأي نسخة عن هذه البيانات السرية أو أي بيانات تستند إلى البيانات الأصلية أو إرسال نسخ عنها إلى أشخاص غير أولئك المذكورين في هذا الاتفاق، إلا بموجب ترخيص رسمي من مكتب الإحصاءات الوطني.

3. تستخدم البيانات حصراً لغايات إحصائية أو في إطار أعمال بحث علمي. كما أنها تهدف فقط إلى تشكيل بيانات مجمعة، بما في ذلك تصميم المعلومات. أما البحث عن أشخاص أو منظمات بحد ذاتها فمحظور بجميع أشكاله. ولا تستخدم البيانات، بأي شكل من الأشكال، لغايات إدارية أو لمصالح تتعلق بالملكية أو لتنفيذ القانون.

4. يتعهد الباحث الرئيسي بالأبدا يحاول أي شخص إعادة تحديد أحد الأشخاص، أو العائلات، أو الشركات أو المنظمات، وعلاوة على ذلك، يمنع استخدام هوية أي شخص أو منظمة تم اكتشافها مصادفةً. وينبغي إبلاغ مركز الإحصاءات الوطني بأي اكتشاف من هذا النوع. كما لا ينبغي أن يعلم بالاكشاف أي شخص لا يرد اسمه في هذا العقد الذي يدير النفاذ إلى البيانات.

5. يتخذ الباحث الرئيسي إجراءات الأمن اللازمة لتجنب أي نفاذ غير مرخص إلى البيانات الجزئية الخاضعة لترخيص والتي قَدَّمها مكتب الإحصاءات الوطني. ويجب أن تلتف البيانات الجزئية تماماً عند انتهاء أعمال البحث، إلا في حال قَدَّم إلى مكتب الإحصاءات الوطني الضمان الكافي لحماية البيانات، وفي حال وافق مكتب الإحصاءات الوطنية خطياً على حفظها. أما إتلاف البيانات الجزئية فيؤكده خطياً الباحث الرئيسي لمكتب الإحصاءات الوطني.

6. يجب أن تحتوي كل الكتب، أو المقالات، أو وثائق العمل، أو الدراسات، أو الأطروحات، أو التقارير، أو غيرها من المنشورات المستندة إلى البيانات التي قَدَّمها مكتب الإحصاءات الوطني على مراجع لمصدرها، وفقاً لطلب ذكر مراجع الاقتباسات الخاص بمجموع البيانات المقدمة.

7. تقدّم إلى مكتب الإحصاءات الوطني نسخة بصيغة إلكترونية تضم كل المنشورات المستندة إلى البيانات المطلوبة.

8. لا تتحمل الجهة التي جمعت البيانات في المرحلة الأولى، ولا مكتب الإحصاءات الوطني، ولا الجهات الممولة المعنية بأي شكل من الأشكال المسؤولية المرتبطة باستخدام البيانات المقدمة أو تحليلها، أو النتائج المستخلصة منها.

9. يسري هذا العقد اعتباراً من تاريخ منح حقوق النفاذ إلى مجموع البيانات السرية ويبقى نافذاً حتى تاريخ نهاية المشروع، أو حتى تاريخ يسبقه إذا انتهى المشروع في وقت مبكر.

10. يتعيّن على الباحث الرئيسي أن يحصل على موافقة مكتب الإحصاءات الوطني المسبقة في ما يتعلق بكل تغيير في خصائص المشروع، وتدابير الأمن، ولكل تغيير يطال فريق العمل أو المنظمة المذكورة في نموذج طلب النفاذ هذا. ويستدعي أي تغيير في قلب المنظمة التي توظف الباحث الرئيسي تقديم طلب جديد، كما تنهي المشروع الأصلي.

11. لا يتساهل مكتب الإحصاءات الوطني في أي انتهاك لهذا العقد، وسيقدم على ملاحظة الأشخاص المسؤولين عن الانتهاك سواء أكان متعمداً أم لا. ويعدّ عدم احترام تعليمات مكتب الإحصاءات الوطني انتهاكاً كبيراً لهذا العقد، وقد يؤدي إلى ملاحقة قضائية. كما سيملك مكتب الإحصاءات الوطني ويشارك مع أرشيفات البيانات الشريكة له قائمة بالأشخاص والمنظمات المسؤولة عن ارتكاب انتهاكات لعقد النفاذ إلى البيانات، وسيقوم هؤلاء الشركاء بمنع بث البيانات في المستقبل لهذه الجهات.

الطرفان الموقعان

قام الباحث الرئيسي أو ممثل مؤهل عن الجهة المستفيدة بقراءة أحكام اتفاق النفاذ إلى البيانات الواردة في القسم و أعلاه والموافقة عليه:

الاسم _____
التوقيع _____
التاريخ _____

قام بدراسة الطلب في (التاريخ)
قرار اللجنة:

قبول الطلب
 رفض الطلب (السبب) _____
 طلب استكمال معلومات _____

الملحق 2

نموذج لسياسة النفاذ إلى مركز بيانات محمي

فيما يلي اقتراح لصياغة السياسة، ينبغي تعديله وفقاً للبلد المعني.

الأهداف

أنشأ أرشيف البيانات الوطن مركز البيانات الوطني المحمي (المشار إليه في ما يلي باسم "المركز") ليسمح للباحثين الذين يتمتعون بمؤهلات معينة بالنفاذ إلى ملفات البيانات الإحصائية السرية، تحت مراقبة مشددة. ويقدم المركز آلية تسمح للباحثين بالاطلاع على ملفات البيانات المفصلة بأمان تام، من دون تعريض مجهولية هوية المستجيبين للخطر.

المكان

يقع المركز في (عنوان المركز، الهاتف، الفاكس، البريد الإلكتروني والموقع الإلكتروني)

أنشطة المركز

يستطيع الباحثون النفاذ إلى البيانات في المركز بإشراف فريق عمل المركز. ويوضع في تصرفهم جهاز حاسوب وبرمجيات، ومكتب.

البيانات

- يعدّ فريق عمل المركز ملفات البيانات اللازمة قبل وصول الباحث ويحرص على ألا تترك المركز أي معلومة سرية.
- بالنسبة إلى الباحثين الذين يودّون إجراء تحليلات عديدة بالاستعانة بمجموعات متعددة من البيانات، سيتمكنون من النفاذ إلى كل مجموعة من البيانات على حدى. ولا يسمح للباحثين بأي حال من الأحوال بدمج مجموعات البيانات بأنفسهم.
- يسمح المركز للباحثين بإحضار بياناتهم الخاصة المغفلة لدراساتها مع مجموعات بيانات المركز وإنشاء قواعد بيانات مدمجة معدة للتخزين في المركز. ويجب أن تكون البيانات التي يقدمها الباحثون بياناتهم الخاصة التي قاموا بجمعها والتي "يملكونها"، أو غيرها من البيانات العامة التي حصلوا عليها بطريقة قانونية. ويتعيّن على الباحثين إلزامياً أن يقدموا إلى المركز توثيقاً كاملاً لكل البيانات التي ستدمج مع بيانات المركز. كما يتحتّم على الباحثين الذين يودّون اللجوء إلى ملفات مدمجة أن يطلبوا مساعدة فريق عمل المركز ليتأكدوا من أنهم يستطيعون دمج بياناتهم ببيانات المركز. ويقبل المركز ملفات البيانات بصيغ SAS، SPSS، و Stata.
- يصدر المركز بشكل دوري نسخاً لكل الملفات على الحاسوب ويحتفظ بها. وتخزّن ملفات الحفظ في مكان آمن، يحصر النفاذ إليه بفريق عمل المركز. كما يمكن أن توضع هذه النسخ في تصرف الباحثين الذين يحتاجون إلى إجراء تحليلات مكتملة. وتحتوي ملفات الحفظ هذه على بيانات قدمها الباحثون وملفات مدمجة. وتتلف هذه الملفات بناءً على طلب خطي من المستخدم.

تجهيزات العمل

- يحتوي المركز على (N) مكاتب لعمل المستخدمين وطابعة بالأسود والأبيض موجودة في غرفة محمية. أما أجهزة الحاسوب في المركز فليست متصلة بشبكة الإنترنت، ومجهزة بحث لا يتمكن المستخدم من إدخال أي واسطة متحركة (CD-ROM، DVD، الأقراص المرنة، مفاتيح USB).
- جهّزت أمكنة العمل في المركز بمعالج (Pentium X XXX MHz) وتشغلت بأنظمة (Windows NT/ غيره).

البرمجيات

- حملت برمجيات CSPro، و EPI-Info، و SAS، و SPS، و Stata على أجهزة الحاسوب المعدة لاستقبال الباحثين إلى جانب حزمة MS Office. ويمكن أن تقدّم لغات تحليل وبرمجة إضافية عند الطلب. وللمزيد من المعلومات بشأن نسخ البرمجيات المتوفرة في المركز، نرجو منكم الاتصال بنا.
- يجب أن يتمتع الباحثون بالمهارات اللازمة لإجراء تحليلاتهم الخاصة بواسطة إحدى التطبيقات البرمجية المقدّمة. إذ لا يقدم المركز خدمة مساعدة تقنية لهذه البرمجيات.

مكان العمل

- يعمل الباحثون تحت إشراف فريق عمل المركز وفي ساعات دوام المكتب الاعتيادية فقط (من الإثنين إلى الجمعة، من الثامنة والنصف صباحاً حتى الخامسة مساءً).
- يحصر النفاذ إلى المركز بالباحثين المذكورين في نموذج تقديم الطلب. وتطلب منهم بطاقة هوية تظهر صورتهم عند الدخول.
- يمكن لثلاثة باحثين يعملون على المشروع نفسه كحد أقصى أن يشغلوا جهاز عمل واحد.
- يطبق المركز قاعدة "الأولوية لمن يصل أولاً".

إشراف فريق عمل المركز (في إطار ضبط الكشف عن السرية)

- مهتّع الباحثون الخارجيون عن إحضار أي وثيقة، أو كتيب، أو دراسة، إلخ، يمكن أن تسمح بالتعرّف إلى هوية المستجيبين أو تكشف عن معلومات شخصية وضعت في تصرف المركز. كما تمنع أجهزة الهاتف، وأجهزة المناداة، وغيرها من وسائل الاتصال مع الخارج.
- لا يسجل الباحثون أي نتيجة، أو ملف، أو برنامج على أي وسيطة بيانات إلكترونية قابلة للنقل. بل يقدم فريق عمل المركز هذه الخدمة عند الطلب.
- لا يحق للباحثين أن يأخذوا نتائج تحليلاتهم إلا بعد لا يحق للباحثين أن يأخذوا نتائج تحليلاتهم إلا بعد يخضعها فريق عمل المركز لفحص ضبط الكشف عن السرية، وتقضي فحوصات ضبط الكشف عن السرية بالبحث عن الخلايا المتعادلة إلى أقل من خمس حالات، والجداول التي تضم

وتدفع المبالغ إلى: (معلومات عن طريقة الدفع)

تقديم مشاريع البحث

يستخدم الباحثون الاستثمار الموجودة في الصفحة التالية لتقديم مشاريعهم. وهو مدعوون إلى التواصل مع فريق عمل المركز للتحقق من أن البيانات التي تهمهم موجودة متوفرة فعلاً، وذلك قبل البدء بصياغة مشروعهم. أما هذا الأخير، فينبغي أن يوصف بحيث يساعد فريق المركز في إعداد الملفات التحليلية اللازمة للمشروع. كما يجب تحديد المتغيرات الضرورية وتحديد الحالات اللازمة في نهاية المطاف بوضوح. ولا تحتوي ملفات البيانات التحليلية إلا على عناصر البيانات اللازمة للتحليلات المنتظرة. وفي هذا الإطار، يتعين على مقدم الطلب أن يحدد في أي أمر يحتاج البيانات التي يود الحصول عليها. أما المشاريع البالغة الضخامة والتعقيد، أو على العكس تلك الهزيلة للغاية، فتتطلب تواصلاً معمقاً بين فريق عمل المركز ومقدمي الطلب، الأمر الذي قد يبطئ العملية. ويمكن أن يتم عمل التحضير لملفات البيانات في مهل معقولة إذا كانت المشاريع الكبيرة والمعقدة مقسمة إلى أجزاء عدة، وإذا كانت البيانات اللازمة لها محددة بوضوح. ويتعين على الباحثين الذين يودون دمج بيانات من المركز بيانات خارجية أن يقدموا هذه الأخيرة إلى فريق عمل المركز قبل القدوم.

يدرس المشروع عند استلامه، وتقيمه لجنة ضبط الكشف عن السرية لهذه الغاية.

ولا بدّ من تطبيق الخصائص الثلاث التالية في فحص المشروع:

- الجدوى التقنية والعلمية للمشروع
- جهوزية موارد المركز
- خطر الكشف عن المعلومات السرية

كما يحيط الباحثون علماً بأن قبول المركز لطلبهم، لا يعني موافقته أو تأييده العام لمنهجية مشروعهم المقترح، أو فرضياته أو سياسته، ولا يعترف له بقيمة مميزة. إذ تمثل موافقة المركز مجرد تقدير لقانونية استخدام ملفا البيانات في إطار الأبحاث الموصوفة، ويشير إلى إمكانية استكمال المشروع بنجاح في المركز.

متغيرات جغرافية، والنماذج التي تضم متغيرات جغرافية (أو متغيرات توازي المتغيرات الجغرافية)، أو حتى قوائم الحالات.

- يجب أن تطبع جميع السجلات أو تُؤرشف بصيغة إلكترونية. يقوم المركز بحفظ هذه الملفات، ولا يحتفظ إلا بالبرامج والعمليات التي يجريها الباحثون الخارجيون. أما التسجيلات التي تخص أبحاثهم الخاصة فلا يتم الاحتفاظ بها.
- تخضع كل النتائج التي يتم الحصول عليها بواسطة برامج الإحصاءات إلى جانب كل المذكرات الخطية المتعلقة بها إلى ضبط للكشف عن السرية يجريه فريق عمل المركز قبل مغادرة المركز. ويمكن نقل الجداول المركبة فقط. ولا يمكن أن يحتوي أي جدول بأي شكل من الأشكال خلايا تتعلق بأقل من خمس حالات تم تحديدها. وعند الاقتضاء، تحذف هذه الخلايا عادة عن طريق المحو. وحرصاً على ألا تتم إعادة بنائها عن طريق خلايا أخرى في الخط أو العمود نفسه، يحذف فريق العمل كل الخطوط والأعمدة الخاصة بهذه الخلايا. وبعد أن تنتهي إجراءات ضبط الكشف عن السرية، يحصل الباحثون على نسخة من الجداول النهائية. ويطبق فريق عمل المركز أفضل الممارسات النافذة لتحديد ما إن كانت البيانات المجدولة تتيح التعرّف إلى هوية المستجيبين أم لا، ويتخذ قرار حريصة استناداً إلى ذلك. وتعدّ قرارات المركز نهائية ولا يمكن للباحثين المفاوضة بشأنها.

رسوم القبول

يتعين على مستخدمي المركز دفع الرسوم العائدة إلى استخدام المكان والتجهيزات، وإشراف فريق العمل، وإجراءات ضبط الكشف عن السرية، وصيانة التجهيزات المعلوماتية (من آليات وبرمجيات)، وإنشاء ملفات البيانات التي يطلبها الباحثون وإدارتها. وتشمل رسوم النفاذ إلى المركز:

تكاليف استخدام التجهيزات (باليوم وحسب عدد أمكنة العمل)	تكاليف إعداد الملفات وإنشائها (التكاليف الثابتة)	الجهة التي ينتمي إليها الباحث الرئيسي
مستخدمون الوطنيون		
مجاناً	مجاناً	فريق عمل وكالة منتمية إلى مركز الإحصاءات الوطنية
(التكلفة / العملة)	(التكلفة / العملة)	وكالات عامة أخرى
(التكلفة / العملة)	(التكلفة / العملة)	الجامعات / مراكز الأبحاث
(التكلفة / العملة)	(التكلفة / العملة)	
مستخدمون الدوليون		
(التكلفة / العملة)	(التكلفة / العملة)	بحث بالشراكة مع مركز الإحصاءات الوطني
(التكلفة / العملة)	(التكلفة / العملة)	المنظمات الدولية
(التكلفة / العملة)	(التكلفة / العملة)	الجامعات / مراكز الأبحاث
(التكلفة / العملة)	(التكلفة / العملة)	المنظمات غير الحكومية

قد يُفرض مبلغ إضافي في حالات العمليات الخاصة، مثل دمج البيانات الإضافية، أو إنشاء صيغ ملفات مصممة خصيصاً، أو حتى الحصول على برمجية مميزة غير معيارية وتحميلها. ويتم الاتفاق على هذا المبلغ بين الباحث وفريق عمل المركز. أما الدفع فيتم سلفاً ويسبق استخدام المركز.

الملحق 3

طلب نفاذ إلى مجموع البيانات الخاضعة لترخيص في إطار مشروع بحث محدد

يجب تكييف نموذج الطلب هذا مع المتطلبات الخاصة المطبقة.

ستبقى المعلومات التي تقدّمونها في نموذج الطلب هذا سرية، إلا في حال انتهاك الاتفاق القانوني المبرم، إذ يبلغ مكتب الاستقصاءات الوطني في هذه الحالة معاهد الإحصاءات الشريكة القائمة في بلدان أخرى بالأمر.

الرجاء إرسال نموذج الطلب هذا عن طريق البريد أو الفاكس إلى العنوان أو الرقم التالي، مرفقاً برسالة من المؤسسة المسؤولة:

العنوان: [العنوان]

الفاكس: [رقم الفاكس]

البريد الإلكتروني (نسخة مصورة): [عنوان البريد الإلكتروني]

العنوان والرقم الخاصين بمجموع (أو مجموعات) البيانات المطلوبة (العنوان الكامل، والسنة، ورقم المرجع كما في فهرس الاستقصاء):

الشروط المطبقة

في هذا الاتفاق،

1. يمثل "الباحث الرئيسي" الباحث الملكّف الذي سيتم التواصل معه بشكل رئيسي في ما يخص هذا الاتفاق. ويتحمل الباحث الرئيسي مسؤولية احترام شروط هذا الاتفاق في ما يتعلق بالنفاذ إلى البيانات. وينبغي أن يكون الباحث الرئيسي مؤهلاً لتمثيل الجهة المستفيدة في إطار هذا النوع من الاتفاقات.

2. يمثل "الباحثون الآخرون" الأشخاص غير الباحث الرئيسي، ومن بينهم المساعدون، الذين سيتمكنون من النفاذ إلى البيانات السرية.

3. تمثل "الجهة المستفيدة" المنظمة/ الجامعة/ المؤسسة التي توظف الباحث الرئيسي.

4. يمثل "ممثل الجهة المستفيدة" الشخص المؤهل لتمثيل الجهة المستفيدة في إطار هذا الاتفاق.

القسم أ. الباحث الرئيسي

- الاسم
- الكنية
- المنصب
- المنظمة
- الوظيفة في المنظمة
- العنوان البريدي

- الهاتف (مع رمز البلد)
- الفاكس (مع رقم البلد)
- البريد الإلكتروني

القسم ب. الباحثون الآخرون

الرجاء ذكر أسماء كل الأعضاء الآخرين في فريق الباحثين المخوّل النفاذ إلى البيانات السرية، ومناصبهم والمنظمات التي ينتمون إليها.

- الإسم والكنية
- الوظيفة
- المنظمة الحاضنة

يرجى إرفاق هذا الطلب بسيرة ذاتية خاصة بكل من يشارك في الأبحاث، مع تحديد جنسياتهم.

القسم ج. الجهة المستفيدة

اسم المنظمة

نوع المنظمة (اختر الإجابة المناسبة)

- وزارة/ إدارة عامة
- جامعة
- مركز أبحاث
- مؤسسة خاصة
- منظمة دولية
- منظمة غير حكومية (وطنية)
- منظمة غير حكومية (دولية)
- غير ذلك (يرجى التحديد)

موقع المنظمة الإلكتروني (رابط URL)

العنوان البريدي

القسم د. ممثل الجهة المستفيدة

- الاسم
- الكنية
- المنصب
- المنظمة
- الوظيفة في المنظمة
- العنوان البريدي
- الهاتف (مع رمز البلد)
- الفاكس (مع رقم البلد)
- البريد الإلكتروني

البرمجية سارية المفعول. ويقوم فريق عمل المركز بتحميل البرمجية لفترة عمل الباحثين (ويبقى الباحث محتفظاً بملكية ترخيص الاستخدام). وفي هذا السياق، نرجو منكم التواصل مع المركز قبل استكمال طلبكم، من أجل التأكد من إمكانية تحقيق المشروع تقنياً.

القسم ح. اتفاق النفاذ إلى البيانات

يبرم الاتفاق التالي، بشرط موافقة الطرفين:

يوافق الباحث الرئيسي والباحثون الآخرون وممثل الجهة المستفيدة على الامتثال إلى الشروط التالية:

1. يحصر الاطلاع على البيانات السرية بالباحث الرئيسي والباحثين الآخرين المحددين في نموذج استمارة الطلب والذين يوقعون تعهداً بالسرية.
2. تستخدم البيانات حصراً لغايات إحصائية. كما أنها تهدف فقط إلى تشكيل بيانات مجمعة، بما في ذلك تصميم المعلومات. أما البحث عن أشخاص أو منظمات بحد ذاتها فمحظور بجميع أشكاله. ولا تستخدم البيانات، بأي شكل من الأشكال، لغايات إدارية أو لمصالح تتعلق بالملكية أو لتنفيذ القانون.
3. يتعهد الباحث الرئيسي ألا يحاول أي شخص إعادة تحديد أحد الأشخاص، أو العائلات، أو الشركات أو المنظمات. وعلاوة على ذلك، يمنع استخدام هوية أي شخص أو منظمة تم اكتشافها مصادفةً. وينبغي إبلاغ المركز بأي اكتشاف من هذا النوع. كما لا ينبغي أن يعلم بالاكتشاف أي شخص لا يرد اسمه في هذا العقد الذي يدير النفاذ إلى البيانات.
4. يجب أن تحتوي كل الكتب، أو المقالات، أو وثائق العمل، أو الدراسات، أو الأطروحات، أو التقارير، أو غيرها من المنشورات المستندة إلى البيانات التي قُدِّمها المركز على مراجع لمصدرها، وفقاً لطلب ذكر مراجع الاقتباسات الخاص بمجموع البيانات المقدمّة.
5. تقدّم إلى المركز نسخة بصيغة إلكترونية تضم كل المنشورات المستندة إلى البيانات المطلوبة.
6. لا تتحمل الجهة التي جمعت البيانات أساساً، ولا المركز، ولا الجهات الممولة المعنية بأي شكل من الأشكال المسؤولية المرتبطة باستخدام البيانات المقدمّة أو تحليلها، أو النتائج المستخلصة منها.
7. لا يتساهل المركز في أي انتهاك لهذا العقد، وسيقدم على ملاحقة الأشخاص المسؤولين عن الانتهاك سواء أكان متعمداً أم لا. ويعدّ عدم احترام تعليمات المركز انتهاكاً كبيراً لهذا العقد، وقد يؤدي إلى ملاحقة قضائية. كما سيملك المركز ويشارك مع أرشيفات البيانات الشريكة له قائمة بالأشخاص والمنظمات المسؤولة عن ارتكاب الانتهاكات لعقد النفاذ إلى البيانات، وسيقوم هؤلاء الشركاء بمنع بث البيانات في المستقبل لهذه الجهات.
8. يحتفظ المركز بحق إيقاف أي مشروع، وفي أي وقت من الأوقات، إذا اعتبر أن أنشطة الباحث لا تراعي السرية أو المعايير الأخلاقية التي ينبغي احترامها في إطار البحث.
9. لا يمكن إخراج أي وثيقة مطبوعة، أو ملف إلكتروني، أو أي ملف أو وسيطة من المركز من دون أن يفحصه مسبقاً فريق عمل المركز لإبعاد أي خطر.
10. يمكن أن تغلق أبواب المركز في وجه الباحث الرئيسي والباحثين الآخرين نهائياً للفترة التي يقدرها المدير ضرورية، وذلك لحماية نزاهة المركز وسريته.

الطرفان الموقعان

قام الطرفان الموقعان أدناه بقراءة أحكام اتفاق النفاذ إلى البيانات الواردة في القسم ح أعلاه والموافقة عليه:

القسم هـ. التعريف بالاستخدام المنتظر للبيانات

يرجى وصف مشروع البحث (الأسئلة، الأهداف، الوسائل، النتائج المتوقعة، الشركاء). لماذا لا يستجيب مجموع البيانات العامة إلى حاجاتكم بالكامل؟ وإذا كانت المعلومات المقدّمة غير كافية، يمكن أن يرفض طلبكم أو أن تتلقوا طلباً بمعلومات مكتملة. ويمكن أن ترفق هذه المعلومات في ملحق لهذا الطلب.

قائمة النتيجة المنتظرة (النتائج المنتظرة) وسياسة البث

هل تتوون بث مجموع البيانات إلى جانب بيانات أخرى؟ نعم لا

إذا كانت الإجابة نعم، يرجى تحديد مجموعات البيانات الأخرى التي ينبغي دمجها.

القسم و. تحديد البيانات والمتغيرات

يقدم المركز على مواقعه الإلكتروني بيانات وصفية مفصلة. وهي تضم وصفاً وملفات البيانات والمتغيرات التي تكوّن كل مجموعة من البيانات. ويرجى من الباحثين تحديد المجموعة الفرعية للمتغيرات أو الحالات التي تهمهم حتى يتمكن المركز من تحضير ملفات البيانات.

يرمي الطلب إلى النفاذ إلى:

- مجموع البيانات الكاملة (كل الملفات والحالات).
- مجموعة فرعية من المتغيرات و/أو الحالات، بما يتوافق مع ما يلي (ملاحظة: تحتوي المجموعات الفرعية بشكل منهجي على المتغيرات مثل معاملات الترجيح الخاصة بالعينة، ومعرّفات التسجيلات).

القسم ز. البرمجيات اللازمة

سيستخدم الباحثون البرمجية التالية:

- CPro
- Stata
- SPSS
- SAS
- برمجية أخرى (يرجى التحديد): _____

ملاحظات:

- يقود المركز بتحديث برمجياته دورياً. لذا يرجى التواصل معنا للمزيد من المعلومات بشأن النسخة المتوفرة لكل تطبيق.
- أما الباحثون الذين يستخدمون برمجية غير مدرجة في قائمة البرمجيات المعيارية التي يقدمها المركز، فيتعيّن عليهم تقديم رخصة لاستخدام هذه

الباحث الرئيسي

الاسم _____
التوقيع _____ التاريخ _____

ممثّل الجهة المستفيدة

الاسم _____
التوقيع _____ التاريخ _____

ينتظر المركز من كل الباحثين أن يلتزموا بالمعايير والمبادئ المتعلقة بالبحث الإحصائي في إطار أعمالها، ولا سبيل لإجراء تحليلات غير تلك التي يتم الموافقة عليها، أما عدم احترام التعليمات، فيترب عليه إلغاء مشروع البحث وحظر دخول المركز لاحقاً.

- dentiality and Microdata Access: Principles
-rodata.access.pdf .Managing Statistical Confi
/Managing.statistical.confidentiality.and.mic
http://www.unece.org/stats/publications
[25] لجنة الأمم المتحدة الاقتصادية لأوروبا (UNECE). 2009. pdf
/Confidentiality_aspects_data_integration
http://www.unece.org/stats/publications
for Statistical or Related Research Purposes.
-ity Aspects of Data Integration Undertaken
Principles and Guidelines on Confidential
[26] لجنة الأمم المتحدة الاقتصادية لأوروبا (UNECE) وإحصاءات السويد.
2003. سرية البيانات الإحصائية والبيانات الجزئية. المناقشات
التي دارت في الندوة لمؤتمر الإحصائيين الأوروبيين في عام 2003.
/publications/statistical.confidentiality.pdf
http://www.unece.org/stats
[27] اللجنة الإحصائية للأمم المتحدة (UNSD). 1994. المبدأ السادس،
/FP-French.htm. المبادئ الأساسية للإحصاءات الرسمية.
http://unstats.un.org/unsd/methods/statorg
[28] مكتب الإحصاء الأمريكي، فرع إدارة المعايير والبرمجيات، قسم دعم
Statistical Methodology Metadata. 2008. الأنظمة.
Survey Design and .واشنغتن دي سي، القسم 3.4.4.
on Statistical Disclosure Limitation Methodology [29]
Working Paper 22 (Revised 2005) – Report
-istical Methodology. 2005. Statistical Policy
/spwp22.html .US Federal Committee on Stat
http://www.fcs.m.gov/working-papers
[30] واتكنز، و. وبويكو، إ. 1996. and Academic Freedom
Data Liberation .معلومات حكومية في كندا 2-3.
/library/gic/v3n2/watkins2/watkins2.html
http://www.usask.ca
[31] واتكنز، و. Canadian Social Science Federation
Model. Unpublished paper written for the
Liberation Initiative: A New Cooperative
/gic/v1n2/watkins/watkins.html .The Data
http://library2.usask.ca

المواقع الإلكترونية

الرابطة الأفريقية لأرشفة البيانات الإحصائية (AASDA)
http://www.aasda.net/home_dev/index.php

مكتب الإحصاءات الأسترالي (ABS)
<http://www.abs.gov.au/>

الجمعية الأمريكية الإحصائية
<http://www.amstat.org/comm/cmtepc/index.cfm?fuseaction=main>

مكتب الإحصاءات المركزي، أيرلندا (CSO)
<http://www.cso.ie/>

المنظمة الأوروبية لأرشفة البيانات في العلوم الاجتماعية (CESSDA)
<http://www.cessda.org/>

Data.Gov (بريطانيا)
<http://data.gov.uk>

Data.Gov (الولايات المتحدة)
<http://www.data.gov/>

مبادرة إتاحة البيانات IDD (باستضافة جامعة أونتاريو - شبكة خدمة البيانات الاقتصادية والاجتماعية)
<https://ospace.scholarsportal.info/handle/1873/69>

تحالف مبادرة توثيق البيانات DDI
<http://www.IDDalliance.org>

قسم التعداد والإحصاء، سري لانكا
<http://statistics.sltidc.lk>

معياري دبلن كور Dublin Core للبيانات الوصفية (DCMI)
<http://dublincore.org/>

خدمة البيانات الاقتصادية والاجتماعية (ESDS)
<http://www.esds.ac.uk/aandp/create/research.asp>

الاستقصاء الاجتماعي الأوروبي
http://www.europeansocialsurvey.org/index.php?option=com_content&task=view&id=78&Itemid=190

المكتب الإحصائي للاتحاد الأوروبي (EUROSTAT)
<http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>

معهد البحوث الاجتماعية والاقتصادية، لجنة استقصاء الأسر البريطانية
<http://www.iser.essex.ac.uk/ulsc/bhps/>

الشبكة الدولية لاستقصاءات الأسر المعيشية (IHSN)
<http://www.ihsn.org>

اتحاد أبحاث العلوم السياسية والاجتماعية بين الجامعات (ICPSR)
<http://www.icpsr.umich.edu>

دراسة الدخل في لوكمبورغ (LIS)
<http://www.lisproject.org/>

Measure DHS
<http://www.measuredhs.com>

مركز ميشيغان للبحوث وبيانات السكان (MCRDC)
www.isr.umich.edu/src/mcrdc/

المركز الوطني للإحصاءات الصحية (NCHS)، مركز بحوث البيانات (الولايات المتحدة)
<http://www.cdc.gov/nchs>

المركز الوطني لبحوث الرأي (NORC) في جامعة شيكاغو
www.norc.org/DataEnclave

مؤسسة العلوم الوطنية (الولايات المتحدة)
<http://www.nsf.gov/index.jsp>

برنامج مراكز البيانات للبحوث والإحصاءات (CDR) - كندا
www.statcan.gc.ca/rdc-cdr/index-fra.htm

إحصاءات كندا، مبادرة إتاحة البيانات (IDD)
<http://www.statcan.gc.ca/dli-ild/dli-idd-fra.htm>

المركز الديمغرافي لأمريكا اللاتينية والكاريبية (CEPAL)، لجنة الأمم المتحدة الإحصائية
<http://www.cepal.org.ar/software/icepa8c.html>

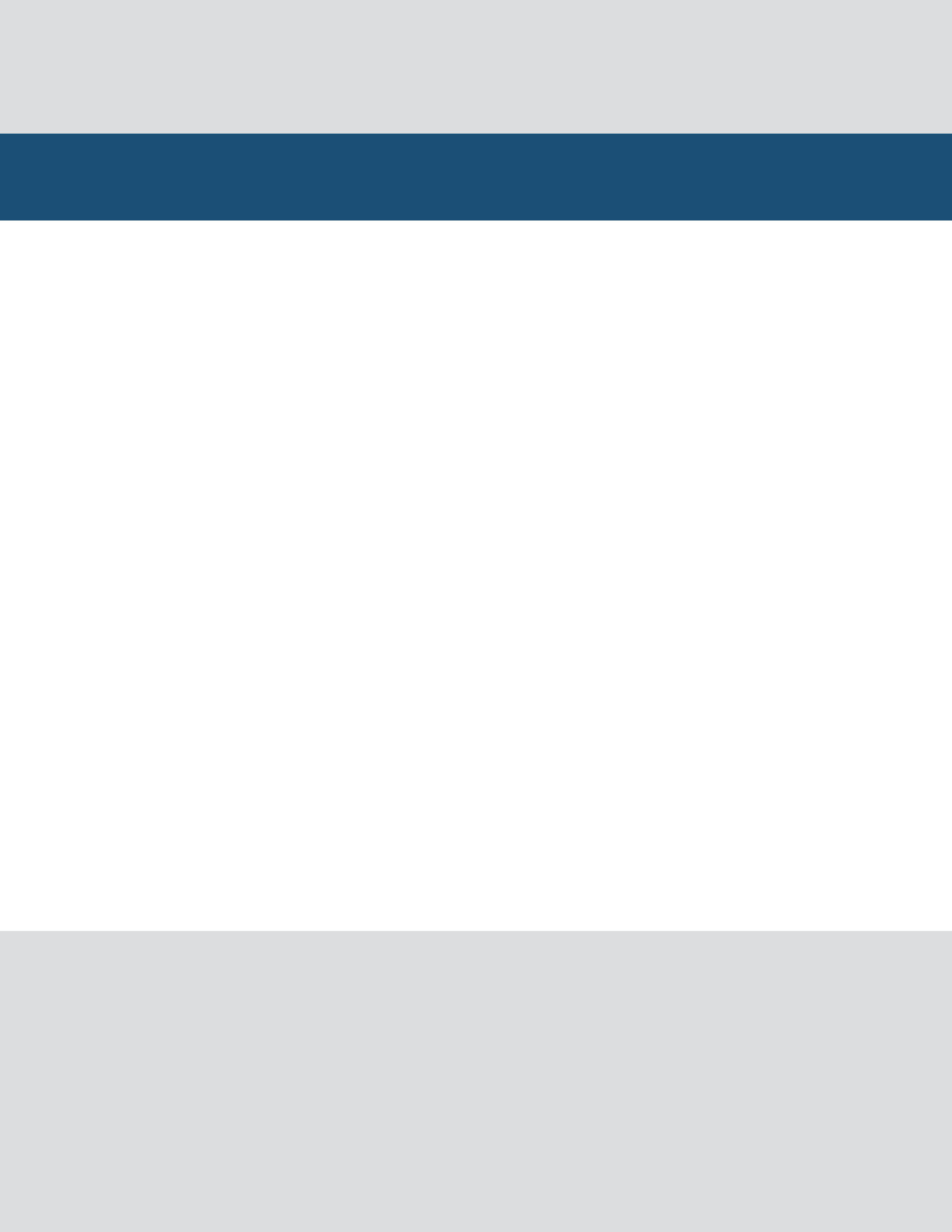
شعبة الإحصاءات في الأمم المتحدة
http://unstats.un.org/unsd/methods/statorg/Principles_stat_activities/principles_stat_activities.asp

أرشيف بيانات المملكة المتحدة (UKDA)
<http://www.dataarchive.ac.uk/sharing/metadata.asp>
<http://securedata.ukda.ac.uk/about/about.asp>

جهاز الإحصاء البريطاني
<http://www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html>

مكتب الإحصاء الأمريكي
<http://www.census.gov/population/www/cen2000/pums/index.html>
www.census.gov/srd/sdc/

موقع ويكيبيديا
<http://en.wikipedia.org>



بخصوص الشبكة الدولية لاستقصاءات الأسر المعيشية IHSN

في فبراير / شباط 2004، التقى ممثلون عن عدد كبير من البلدان ووكالات التنمية في مراكش بالمغرب، في إطار الاجتماع الثاني الدولي للمائدة المستديرة المعني بإدارة نتائج التنمية. وقد تمثلت المهام في دراسة كيفية تحسين عملية التنسيق بين مساعدات الممولين، من أجل دعم النظام الإحصائي وتعزيز قدرات المتابعة والتقييم اللازمة حتى تأخذ الدول على عاتقها إدارة عملية التنمية فيها. وقد نتج عن هذا اللقاء، من ضمن جملة من الأمور، اعتماد خطة عمل مراكش للإحصاء (MAPS).

أما توصيات خطة العمل فكان من أبرزها إنشاء شبكة دولية لاستقصاءات الأسر المعيشية. وقد اعترف المجتمع الدولي من خلال إنشاء هذه الشبكة بالدور الحاسم الذي تضطلع به نماذج الاستقصاءات في تخطيط السياسات وبرامج التنمية وتنفيذها ومتابعتها. كما قُدمت إلى المنظمات الوطنية والدولية منصة تسمح بتحسين عمليتي التنسيق والإدارة فيما يتعلق بجمع البيانات الاجتماعية الاقتصادية وتحليلها، وتتيح حشد الوسائل لتعزيز الفعالية والكفاءة في مجال تحقيق الاستقصاءات في البلدان النامية.

وتهدف سلسلة وثائق عمل الشبكة الدولية لاستقصاءات الأسر المعيشية IHSN إلى إغناء نقاش الأفكار الدائر بشأن تأليف الاستقصاءات وإجرائها لدى الأسر، وتحليل البيانات التي تجمع بواسطة هذه الاستقصاءات وتحليلها وبنائها واستخدامها. أما إذا رغبتكم في تقديم نص بهدف نشره في سلسلة وثائق عمل الشبكة الدولية لاستقصاءات الأسر المعيشية IHSN، فنرجو منكم الاتصال بأمانة الشبكة على العنوان الإلكتروني التالي: info@ihnsn.org.

www.ihnsn.org

العنوان الإلكتروني: info@ihnsn.org

www.ihnsn.org
E-mail: info@ihnsn.org