



الهيئة الاتحادية
للتنافسية والإحصاء
FEDERAL COMPETITIVENESS
AND STATISTICS AUTHORITY



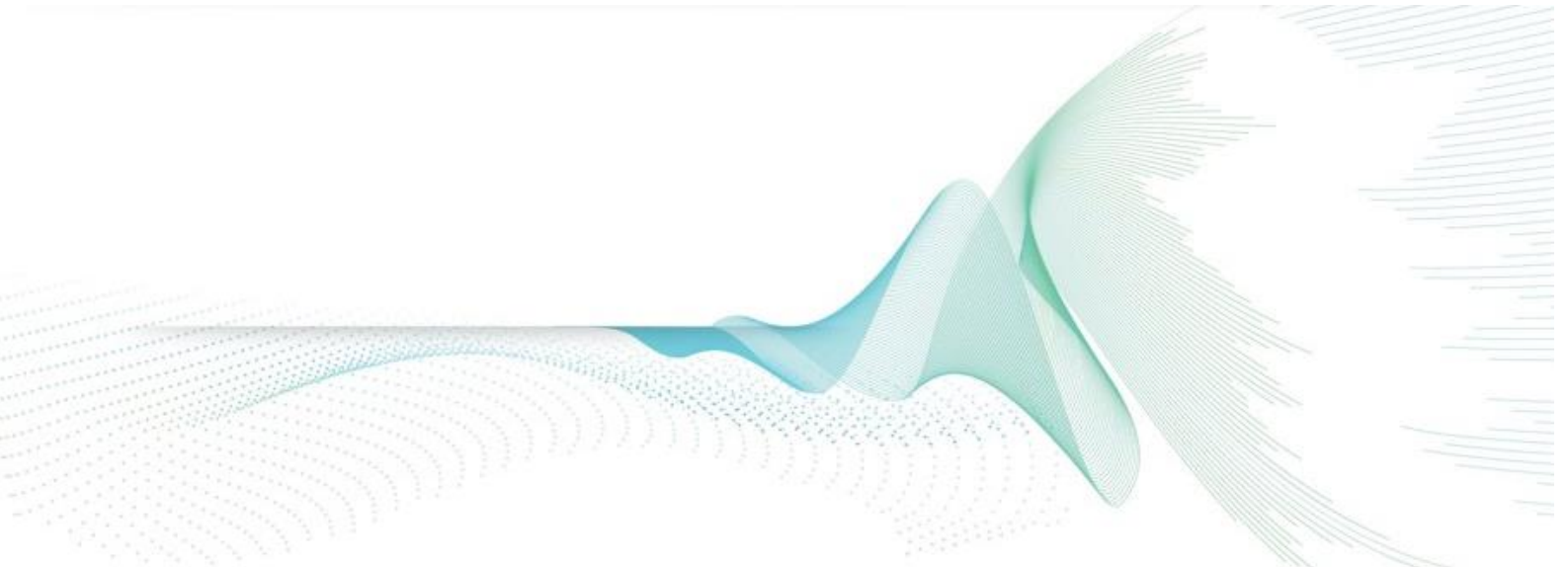
مؤشر نضج البيانات

الدليل الإرشادي لمقاييس ومؤشرات جودة البيانات الإدارية

الأدلة الإرشادية لمؤشر نضج البيانات الحكومية

يوليو 2020

النسخة الأولى 01



فهرس المحتويات

3	1. مقدمة
4	2. محاور مؤشر نضج البيانات الحكومية
5	3. معايير مؤشر نضج البيانات الحكومية
6	4. تطبيق معايير قياس مستوى جودة البيانات
7	1.4 الشمولية (COVERAGE)
8	2.4 الدقة (ACCURACY)
9	3.4 الاتساق (COHERENCE)
10	4.4 الحدائة (TIMELINESS)
10	5.4 إمكانية الوصول (ACCESSIBILITY)
11	6.4 الاكتمال (COMPLETENESS)
11	7.4 التحقق من صحة البيانات (VALIDATION)
12	8.4 التوصيف (DESCRIPTIVENESS)

1. مقدمة

تستند القرارات والخطط التي تتبناها الحكومات إلى بيانات ومعلومات تمتاز بالجودة والحداثة، وتعتبر البيانات الحكومية أو ما يعرف بالبيانات الإدارية أو البيانات السجلية أو السجلات الإدارية، أحد أهم المصادر الأساسية لرسم السياسات، وصنع القرارات، لذا ينبغي أن تمتاز هذه البيانات بمستوى عالٍ من الجودة والدقة والشمول والحداثة.

ينسجم إطلاق مؤشر نضج البيانات الحكومية مع توجهات مكتب رئاسة مجلس الوزراء باعتبار الهيئة الاتحادية للتنافسية والإحصاء هي الجهة الممكنة في مجال نضج البيانات، للارتقاء بتنافسية دولة الإمارات العربية المتحدة، وتوفير إحصاءات ومعلومات دقيقة لمتخذي القرار وراسمي السياسات، وللمخططين والباحثين والإعلاميين ورجال الأعمال وللمؤسسات الوطنية والإقليمية والمنظمات الدولية، مع ضرورة تحقيق الاتساق عند نشر بيانات الدولة، بهدف تحقيق رؤية الإمارات 2021 والأجندة الوطنية في أن تبني دولة الإمارات أفضل اقتصاد تنافسي معرفي.

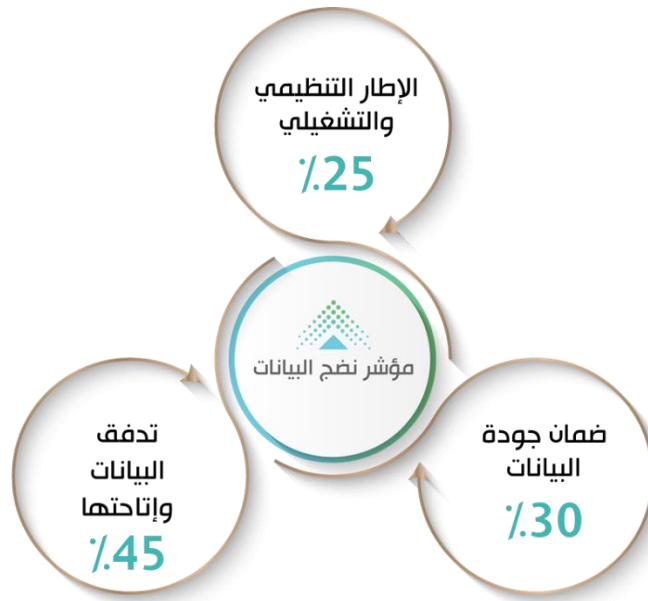
وتسعى الهيئة الاتحادية للتنافسية والإحصاء لتمكين الجهات الاتحادية من تحقيق التميز في إدارة البيانات وتبادلها ورفع وتحسين مستوى جودتها من خلال من خلال توفير دليل إرشادي يحقق مفهوم الجودة الشاملة للسجلات الإدارية وتحديد مجموعة من الإجراءات والضوابط التي تساهم في تحقيق معيار ضمان جودة البيانات في كافة الجهات الاتحادية.

يعرض هذا الدليل أهم المقاييس التي يجب أن تتصف بها البيانات الحكومية وهي جزء سيتم الاستناد عليه في قياس مؤشر نضج البيانات الحكومية لعام 2020.

2. محاور مؤشر نضج البيانات الحكومية

ينقسم المؤشر إلى ثلاثة محاور رئيسية تمثل الدعائم الأساسية لإدارة البيانات في الجهات الحكومية الاتحادية، وتعتمد أوزانها على أهمية كل محور على النحو التالي:

المحور	الوزن
الإطار التنظيمي والتشغيلي	25%
ضمان جودة البيانات	30%
تدفق البيانات وإتاحتها	45%



شكل(1): المحاور الرئيسية لمؤشر نضج البيانات الحكومية

3. معايير مؤشر نضج البيانات الحكومية

تعرف البيانات الحكومية على أنها مجموعة من البيانات التي تنشأ نتيجة لعمليات التسجيل الإداري من خلال الخدمات التي تقدمها الجهات المختلفة، لاسيما الحكومية منها للأفراد والمؤسسات، ونتيجة للإجراءات المختلفة التي تقوم بها الجهات الحكومية تنشأ قواعد بيانات إلكترونية تتعلق بالأفراد والشركات والعاملين وغيرها كل حسب اختصاصه.

فيما يلي جدول بمحاور ومعايير مؤشر نضج البيانات الحكومية:

الوزن	المعايير	الوزن	المحاور الفرعية	الوزن	المحاور الرئيسية
10%	1.1.1 تشكيل فريق إدارة البيانات وتحديد المهام (Data Governance)	10%	1.1 حوكمة الإطار التنظيمي لإدارة البيانات	25%	(1) الإطار التنظيمي والتشغيلي
5%	1.2.1 تعزيز قدرات وإمكانيات الفريق في مجال حوكمة وإدارة البيانات (Skills)	15%	2.1 تطبيق الإطار التشغيلي لإدارة البيانات		
10%	2.2.1 إعداد خارطة الطريق وخطة الامتثال لمتطلبات إطار إدارة البيانات (Roadmap)				
5%	1.1.2 حصر أنظمة وقواعد وأصول البيانات (Info. Systems & Databases)	10%	1.2 حصر مجموعات البيانات وتصنيفها	30%	(2) جودة البيانات الحكومية
5%	2.1.2 مواءمة تصنيفات بيانات الجهة (Compliance)	20%	2.2 ضمان جودة البيانات		
10%	1.2.2 تطبيق معايير قياس مستوى جودة البيانات (Data Quality)				
10%	2.2.2 مواءمة البيانات الخام مع المؤشرات (Raw Data)				
15%	1.1.3 الامتثال لمتطلبات إتاحة البيانات (Data Access Requirements)	25%	1.3 إتاحة البيانات	45%	(3) إتاحة البيانات وتدقيقها
10%	2.1.3 الامتثال لبرنامج البيانات الوطنية (Data sharing Calendar)				
10%	1.2.3 تفعيل الربط الإلكتروني لتبادل البيانات (Interoperability)	20%	2.3 تبادل البيانات		
10%	2.2.3 دورية مشاركة البيانات (Data Sharing Frequency)				

4. تطبيق معايير قياس مستوى جودة البيانات

تساعد مجموعة المعايير الواردة أدناه في تأسيس منظومة متكاملة لضبط جودة بيانات السجلات الإدارية للجهات الاتحادية، حيث تم إعداد هذا الدليل استناداً لما ورد من توصيات وتوجهات في الإطار الوطني لمعايير جودة البيانات الصادر عن الهيئة الاتحادية للتنافسية والإحصاء والأدلة والتي أكدت على أن أي منظومة لضبط جودة البيانات السجلية لابد وأن تضمن العناصر التالية والتي سيتم اعتمادها كمقياس للمؤشر:

(1) الشمولية (Coverage)

(2) الدقة (Accuracy)

(3) الاتساق (Coherence)

(4) الحداثة (Timeliness)

(5) إمكانية الوصول (Accessibility)

(6) الاكتمال (Completeness)

(7) التحقق من صحة البيانات (Validation)

(8) التوصيف (Descriptiveness)

يتكون كل مقياس من مؤشر أو أكثر يمكن قياسه من خلال البيانات التي يتم إتاحتها للهيئة الاتحادية للتنافسية والإحصاء بحيث تعكس مستوى تحقيق المقياس الخاص بالجودة.

يتم تقييم المؤشرات وفقاً لما يلي:

الوصف	القيمة
عدم تحقق متطلبات المؤشر أو تحققه بشكل بسيط (عدم تحقق)	0
تحقق متطلبات المؤشر بشكل متوسط أو جيد (تحقق جزئي)	0.5
تحقق متطلبات المؤشر بشكل كامل (تحقق كلي)	1
في حال كان المؤشر "لا ينطبق" على الجهة، فإن وزن المؤشر ودرجته ستتوزع بشكل تلقائي على بقية المؤشرات في نفس المقياس	لا ينطبق

1.4 الشمولية (Coverage)

في إطار تحسين مستوى شمولية البيانات، ينبغي على الجهات الحكومية الاتحادية توفير تقرير يقيس نسبة تغطية البيانات من إجمالي البيانات المطلوبة لرصد حالات النقص. كما يتم حصر البيانات وتصنيفها ضمن مبادرة معايير البيانات الوطنية الذكية وفقاً للتالي:

الرقم	المؤشر	المفهوم	التقييم
1.1	التغطية الجغرافية	تغطية البيانات لكافة المناطق الجغرافية، على سبيل المثال: - الإمارة والمدن والمناطق الرئيسية والمناطق الفرعية حسب تصنيف الإمارة.	(0) تغطية ضعيفة : توجد إمارة أو أكثر لم يتم تغطيتها ضمن نطاق البيانات (0.5) تغطية جيدة : تم تغطية جميع الإمارات ضمن نطاق البيانات مع وجود نقص في بيانات بعض المناطق الرئيسية والفرعية (1) تغطية كاملة : تم تغطية جميع الإمارات والمناطق الرئيسية والفرعية المطلوبة بشكل كامل
1.2	شمول أو تغطية الفئات المستهدفة	تغطية البيانات لجميع الفئات المستهدفة حسب المتطلبات على سبيل المثال: - الفئات الاقتصادية كالأنشطة الاقتصادية وعدد العاملين وغيرها. - الخصائص السكانية كالنوع والحالة الاجتماعية وغيرها	(0) تغطية ضعيفة : توجد فئة رئيسية أو أكثر لم يتم تغطيتها ضمن نطاق البيانات المطلوبة (0.5) تغطية جيدة : تم تغطية جميع الفئات مع وجود نقص بسيط في بعضها وتأثيرها بسيط على شمولية البيانات (1) تغطية كاملة : تم تغطية جميع الفئات في البيانات بشكل كامل

2.4 الدقة (Accuracy)

في إطار السعي نحو دقة البيانات، ينبغي على الجهات الاتحادية مراقبة ومراجعة البيانات بشكل دوري خلال عملية تجميعها ومعالجتها وتجهيزها لضمان دقتها، ويمكن أن تقاس الدقة من خلال المؤشرات التالية:

الرقم	المؤشر	المفهوم	التقييم
2.1	التكرار	تكرار نفس الحالة في الخصائص على سبيل المثال: - تكرار الخصائص مثل الرقم التعريفي، أو الشركة... إلخ - تكرار الرقم التعريفي أو اسم الشركة مع اختلاف باقي الخصائص أو الحقول	(0) تكرار مؤثر : وجود تكرارات كثيرة في البيانات تؤثر على دقتها بشكل كبير (نسبة التكرار تتعدى 20%) (0.5) تكرار غير مؤثر : وجود تكرارات بسيطة في البيانات تؤثر على دقتها بشكل بسيط (نسبة التكرار من 5% إلى 20%) (1) لا يوجد تكرار : لا يوجد أي تكرار في البيانات (نسبة التكرار أقل من 5%)
2.2	القيم الخاطئة	القيم التي تختلف عن مواصفات البيانات أو تتعارض مع القيم الحقيقية، على سبيل المثال: - وجود نص في خانة رقمية وقد يكون العكس مثل وجود نص أو وصف في خانة العمر القيم التي تكون خارج المدى المتوقع أو المنطق (مثلاً شخص عمره 180 سنة) الحالة الاجتماعية متزوج والعمر 13 سنة	(0) قيم خاطئة مؤثرة : وجود مجموعة كبيرة من القيم الخاطئة والتي تؤثر على دقة البيانات بشكل كبير (نسبة القيم الخاطئة تتعدى 20%) (0.5) قيم خاطئة غير مؤثرة : وجود مجموعة بسيطة من القيم الخاطئة والتي تؤثر على دقة البيانات بشكل بسيط (نسبة القيم الخاطئة من 5% إلى 20%) (1) لا يوجد قيم خاطئة : لا توجد أية قيم خاطئة في البيانات (نسبة القيم الخاطئة أقل من 5%)

3.4 الاتساق (Coherence)

في إطار الاستفادة القصوى من البيانات، ينبغي على الجهات الحكومية فحص اتساق ومنطقية البيانات التي تقوم الجهة بجمعها، من خلال إجراء مجموعة من المقارنات سواء مع بيانات تاريخية (سلسلة زمنية) أو من خلال مقارنة هذه البيانات مع بيانات مماثلة من مصادر أخرى:

الرقم	المؤشر	المفهوم	التقييم
3.1	الاتساق الزمني	مقارنه التغيير المنطقي في البيانات عبر الزمن وعدم وجود تغيرات كبيرة أو غير منطقية أو غير مبررة	<p>(0) <u>بيانات غير متسقة مؤثرة</u>: يوجد حالات كثيرة من عدم الاتساق وتغيرات في البيانات دون تبرير منطقي يحاكي الواقع في البيانات عند مقارنتها مع البيانات التاريخية السابقة، ولها تأثير كبير على جودة البيانات لبعض الخصائص</p> <p>(0.5) <u>بيانات غير متسقة وغير مؤثرة</u>: نسبة التأثير بسيطة بسبب عدم الاتساق في البيانات عند مقارنتها مع البيانات التاريخية السابقة، ولها تأثير بسيط على جودة البيانات لبعض الخصائص</p> <p>(1) <u>بيانات متسقة</u>: لا يوجد حالات عدم اتساق في البيانات عند مقارنتها مع البيانات التاريخية السابقة لجميع الخصائص</p>
3.2	الاتساق مع مصادر أخرى	مقارنة منطقية البيانات الحالية مع بيانات مماثلة من مصادر أخرى سواءً أكانت بيانات صادرة من نفس الجهة أو من جهات حكومية أخرى. ويمكن أن يساعد ربط مجموعات البيانات مع مصادر سجلية رئيسية في تحقيق الاتساق	<p>(0) <u>بيانات غير متسقة مؤثرة</u>: يوجد حالات كثيرة من عدم الاتساق في البيانات عند مقارنتها مع مصادر أخرى، ولها تأثير كبير على جودة البيانات لبعض الخصائص</p> <p>(0.5) <u>بيانات غير متسقة وغير مؤثرة</u>: يوجد حالات بسيطة من عدم الاتساق في البيانات عند مقارنتها مع مصادر أخرى، ولها تأثير بسيط على جودة البيانات لبعض الخصائص</p> <p>(1) <u>بيانات متسقة</u>: يتم الربط الإلكتروني مع الجهة الحكومية الرئيسية لاستكمال كافة الخصائص والفئات/أو لا يوجد حالات عدم اتساق في البيانات عند مقارنتها مع مصادر أخرى لجميع المتغيرات</p>
3.3	الاتساق الجغرافي	منطقية البيانات عند مقارنتها بين الإمارات والمناطق الجغرافية في الدولة، بحيث تكون البيانات متناسقة مع التغيير الجغرافي	<p>(0) <u>بيانات غير متسقة مؤثرة</u>: يوجد حالات كثيرة من عدم الاتساق في البيانات عند مقارنتها جغرافياً ولها تأثير كبير على جودة البيانات لبعض الخصائص</p> <p>(0.5) <u>بيانات غير متسقة وغير مؤثرة</u>: يوجد حالات بسيطة من عدم الاتساق في البيانات عند مقارنتها جغرافياً ولها تأثير بسيط على جودة البيانات لبعض الخصائص</p> <p>(1) <u>بيانات متسقة</u>: لا يوجد حالات / حالات بسيط جداً من عدم اتساق في البيانات عند مقارنتها جغرافياً لجميع الخصائص</p>

4.4 الحدائة (Timeliness)

في إطار الجهود الحكومية نحو إتاحة البيانات، ينبغي أن تسعى الجهات الحكومية نحو إتاحة بياناتها وفقاً لاحتياجات صنّاع القرار، لتعزيز فائدتها في عملية التخطيط ورسم السياسات وصناعة القرار بشكل دوري ومحدث، ويرتبط مفهوم الحدائة بتاريخ مرجعية البيانات

الرقم	المؤشر	المفهوم	التقييم
4.1	مستوى الحدائة	الحدائة هي ضمان توفير البيانات ضمن الفترة الزمنية المرجعية المحددة	<p>(0) بيانات غير حدائة: الفترة المرجعية للبيانات المستلمة أقدم من الفترة المرجعية المحددة للبيانات المطلوبة تتعدى العامين</p> <p>(0.5) بيانات متوسطة الحدائة: الفترة المرجعية للبيانات المستلمة أقدم من الفترة المرجعية المحددة للبيانات المطلوبة من عام إلى عامين</p> <p>(1) بيانات حدائة: البيانات المستلمة تمثل فترة مرجعية لا تتعدى عام واحد</p>

5.4 إمكانية الوصول (Accessibility)

في إطار تعزيز مستوى إتاحة البيانات، ينبغي على الجهات الاتحادية إتاحة البيانات وتسهيل الوصول للبيانات بصيغ تنسيقية ملائمة (Format) وقابلة للتبادل الإلكتروني

الرقم	المؤشر	المفهوم	التقييم
5.1	إمكانية الوصول والتعامل مع البيانات	توفر البيانات ضمن صيغ تنسيقية (Format) متوافقة وملائمة يمكن من خلالها التعامل مع البيانات وقراءتها إلكترونياً دون الحاجة لأدوات أو برمجيات خاصة	<p>(0) بيانات بصيغ تنسيقية غير متوافقة: الصيغ التنسيقية للبيانات المستلمة لا تتوافق مع الصيغ التنسيقية القياسية المعتمدة وفقاً لمتطلبات الهيئة في تبادل البيانات الإلكتروني</p> <p>(0.5) بيانات بصيغ تنسيقية متوافقة بشكل جزئي: الصيغ التنسيقية للبيانات المستلمة تتوافق جزئياً مع الصيغ التنسيقية القياسية المعتمدة وفقاً لمتطلبات الهيئة في تبادل البيانات الإلكتروني</p> <p>(1) بيانات بصيغ تنسيقية متوافقة بالكامل: الصيغ التنسيقية للبيانات المستلمة تتوافق بشكل تام مع الصيغ التنسيقية القياسية المعتمدة وفقاً لمتطلبات الهيئة في تبادل البيانات الإلكتروني</p>

6.4 الاكتمال (Completeness)

لغرض ضمان توفر البيانات بشكل يتلاءم مع متطلبات الاستفادة منها في الوقت الحاضر والمستقبل، ينبغي أن تكون مجموعة البيانات التي يتم مشاركتها متكاملة ويمكن تفسيرها بشكل منطقي وواضح، وأن تكون قابلة للاستخدام دون الحاجة إلى بيانات مرجعية إضافية لتفسيرها أو استخدامها.

الرقم	المؤشر	المفهوم	التقييم
6.1	القيم المفقودة	هي الخانات أو الحقول في مجموعة البيانات التي لا تحتوي على أي قيمة أي أن بعض الخصائص أو المتغيرات لم يتم استكمال تعبئتها	(0) <u>قيم مفقودة مؤثرة</u> : وجود مجموعة كبيرة من القيم المفقودة تؤثر على دقة البيانات بشكل كبير (نسبة القيم المفقودة تتعدى 20%) (0.5) <u>قيم مفقودة غير مؤثرة</u> : وجود قيم بسيطة مفقودة وتؤثر على دقة البيانات بشكل بسيط (نسبة القيم المفقودة من 5% إلى 20%) (1) <u>لا يوجد قيم مفقودة</u> : لا توجد أية قيم مفقودة في البيانات (نسبة القيم المفقودة أقل 5%)

7.4 التحقق من صحة البيانات (Validation)

في إطار ضمان المستوى الأمثل لجودة البيانات قبل مشاركتها من قبل الجهات الاتحادية، ينبغي على الجهات الاتحادية تنفيذ إجراءات التدقيق والتحقق من صحة البيانات خلال عمليات إدخال البيانات أو تبادلها ضمن الأنظمة وقواعد البيانات المصدريّة لدى الجهة الاتحادية.

الرقم	المؤشر	المفهوم	التقييم
7.1	التدقيق والتحقق من صحة البيانات	القيام بتطبيق قواعد التحقق من صحة البيانات (Data Validation Rules) على مستوى الأنظمة الإلكترونية وقواعد البيانات	(0) <u>لا يتم تطبيق قواعد التدقيق والتحقق</u> : لا يوجد أي قواعد للتحقق من صحة البيانات مطبقة على مستوى الأنظمة أو قواعد البيانات لدى الجهة الاتحادية (0.5) <u>يتم تطبيق قواعد التدقيق والتحقق بشكل جزئي</u> : يتم تطبيق بعض قواعد التدقيق والتحقق من صحة البيانات على مستوى الأنظمة أو قواعد البيانات لدى الجهة الاتحادية بشكل جزئي (1) <u>يتم تطبيق قواعد التدقيق والتحقق بشكل كامل</u> : يتم تطبيق قواعد التدقيق والتحقق من صحة البيانات على مستوى الأنظمة أو قواعد البيانات بشكل كامل لدى الجهة الاتحادية

8.4 التوصيف (Descriptiveness)

في إطار تعزيز ورفع مستوى تدفق ومشاركة البيانات الحكومية، ينبغي على الجهات الاتحادية توثيق وإتاحة قابلية الاطلاع على البيانات الوصفية الشاملة لمجموعات البيانات التي يتم مشاركتها مع الهيئة وفقاً للمعيار التالي:

الرقم	المؤشر	المفهوم	التقييم
8.1	إمكانية الاطلاع على البيانات الوصفية	إتاحة البيانات الوصفية (Metadata) من قبل الجهة المزودة للبيانات ومدى سهوله الوصول والاطلاع عليها للمستفيدين من مجموعات البيانات	(0) <u>بيانات وصفية غير متاحة</u> : البيانات الوصفية الخاصة بمجموعات البيانات المستلمة غير متاحة (0.5) <u>بيانات وصفية متاحة بشكل جزئي</u> : البيانات الوصفية الخاصة بمجموعات البيانات المستلمة غير مكتملة / جزئية (1) <u>بيانات وصفية متاحة بالكامل</u> : البيانات الوصفية الخاصة بمجموعات البيانات المستلمة مكتملة ويمكن الوصول إليها



@FCSAUAE | @UAESDGS

