المركز الاتحادي للتنافسية والإحصاء
FEDERAL COMPETITIVENESS AND STATISTICS CENTRE

United Arab Emirates

مؤشر نضج البيانات
DATA MATURITY INDEX

الدورة الثانية 2021

ورشة جودة البيانات الحكومية

يوليو 2021

محاور المؤشر

تم تقسيم المؤشر إلى 3 محاور رئيسية تمثل الدعائم الأساسية لإدارة البيانات في الجهات الاتحادية ووضع أوزان حسب أهمية المحور كالآتي:

(1) الإطار التنظيمي والتشغيلي ٪15

(2) ضمان جودة البيانات ٪35

(3) تدفق البيانات وإتاحتها ٪50

تعريف 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.

الأوزان المحددة للمحاور ومعايير مؤشر نضج البيانات الحكومية

المحاور الرئيسية	وزن المحور الرئيسي	المحاور الفرعية	وزن المحور الفرعي	المعايير	وزن المعيار
(1) الإطار التنظيمي والتشغيلي	٪15	1.1 حوكمة الإطار التنظيمي لإدارة البيانات	٪5	1.1.1 تشكيل فريق إدارة البيانات وتحديد المهام (Data Governance)	٪5
		1.2 تطبيق الإطار التشغيلي لإدارة البيانات	٪10	1.2.1 تعزيز قدرات وإمكانيات الفريق في مجال حوكمة وإدارة البيانات (Skills)	٪5
				1.2.2 إعداد خطة الامتثال لمتطلبات إطار إدارة البيانات (Roadmap)	٪5
(2) جودة البيانات الحكومية	٪35	2.1 حصر مجموعات البيانات وتصنيفها	٪15	2.1.1 حصر أنظمة وقواعد وأصول البيانات (Info. Systems & Databases)	٪5
				2.1.2 مواءمة تصنيفات بيانات الجهة (Compliance)	٪10
		2.2 ضمان جودة البيانات	٪20	2.2.1 تطبيق معايير قياس مستوى جودة البيانات (Data Quality)	٪10
				2.2.2 مواءمة البيانات الخام مع المؤشرات (Raw Data)	٪10
(3) إتاحة البيانات وتدفقها	٪50	3.1 إتاحة البيانات	٪25	3.1.1 الامتثال لمتطلبات إتاحة البيانات (Data Access Requirements)	٪15
				3.1.2 الامتثال لرزنامة البيانات الوطنية (Data sharing Calendar)	٪10
		3.2 تبادل البيانات	٪25	3.2.1 تفعيل الربط الإلكتروني لتبادل البيانات (Interoperability)	٪15
				3.2.2 دورية مشاركة البيانات (Data Sharing Frequency)	٪10

معايير قياس مستوى جودة البيانات

تساعد مجموعة المعايير الخاصة بقياس مستوى جودة البيانات (المقاييس) المذكورة أدناه في تأسيس منظومة متكاملة لضبط جودة بيانات السجلات الإدارية للجهات الاتحادية.

(1) إمكانية الوصول (Accessibility)

(2) الدقة (Accuracy)

(3) التوصيف (Descriptiveness)

(4) الحداثة (Timeliness)

(5) الاكتمال (Completeness)

(6) التحقق من صحة البيانات (Validation)

(7) الشمولية (Coverage)

(8) الاتساق (Coherence)

آلية التقييم لمعايير قياس مستوى جودة البيانات

- يتكون كل مقياس مستوى نضج البيانات من مؤشر واحد أو أكثر
- يتم قياس كل مؤشر من خلال البيانات التي يتم إتاحتها للمركز الاتحادي للتنافسية والإحصاء بحيث تعكس مستوى تحقيق المقياس الخاص بالجودة.
- يتم تحديد قيمة مستوى جودة البيانات لكل مؤشر وفقاً للمستويات التالية :

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

تفاصيل معايير قياس مستوى جودة البيانات

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

(1) إمكانية الوصول (Accessibility)

(2) الدقة (Accuracy)

(3) التوصيف (Descriptiveness)

(4) الحداثة (Timeliness)

(5) الاكتمال (Completeness)

(6) التحقق من صحة البيانات (Validation)

(7) الشمولية (Coverage)

(8) الاتساق (Coherence)

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
إمكانية الوصول والتعامل مع البيانات	توفر البيانات ضمن صيغ تنسيقية (Format) متوافقة وملائمة يمكن من خلالها التعامل مع البيانات وقراءتها إلكترونيا دون الحاجة لأدوات أو برمجيات خاصة	0	بيانات بصيغ تنسيقية غير متوافقة	الصيغ التنسيقية للبيانات المستلمة لا تتوافق مع الصيغ التنسيقية القياسية المعتمدة وفقاً لمتطلبات المركز في تبادل البيانات الإلكتروني
		0.5	بيانات بصيغ تنسيقية متوافقة بشكل جزئي	الصيغ التنسيقية للبيانات المستلمة تتوافق جزئياً مع الصيغ التنسيقية القياسية المعتمدة وفقاً لمتطلبات المركز في تبادل البيانات الإلكتروني
		1	بيانات بصيغ تنسيقية متوافقة بالكامل	الصيغ التنسيقية للبيانات المستلمة تتوافق بشكل تام مع الصيغ التنسيقية القياسية المعتمدة وفقاً لمتطلبات المركز في تبادل البيانات الإلكتروني

تفاصيل معايير قياس مستوى جودة البيانات

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:

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
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 من5٪ إلى 20٪
		1	لا يوجد تكرار	لا يوجد أي تكرار في البيانات (نسبة التكرار أقل من 5٪)

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
القيم الخطأ	القيم التي تختلف عن مواصفات البيانات أو تتعارض مع القيم الحقيقية، على سبيل المثال: • وجود نص في خانات رقمية وقد يكون العكس مثل وجود نص أو وصف في خانات العمر القيم التي تكون خارج المدى المتوقع أو المنطق (مثلاً شخص عمره 180 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٪)

(1) إمكانية الوصول (Accessibility)

(2) الدقة (Accuracy)

(3) التوصيف (Descriptiveness)

(4) الحداثة (Timeliness)

(5) الاكتمال (Completeness)

(6) التحقق من صحة البيانات (Validation)

(7) الشمولية (Coverage)

(8) الاتساق (Coherence)

تفاصيل معايير قياس مستوى جودة البيانات

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 المركز وفقاً للمعيار التالي:

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
إمكانية الاطلاع على البيانات الوصفية	إتاحة البيانات الوصفية (Metadata) من قبل الجهة المزودة للبيانات ومدى سهوله الوصول والاطلاع عليها للمستفيدين من مجموعات البيانات	0	بيانات وصفية غير متاحة	البيانات الوصفية الخاصة بمجموعات البيانات المستلمة غير متاحة
		0.5	بيانات وصفية متاحة بشكل جزئي	البيانات الوصفية الخاصة بمجموعات البيانات المستلمة غير مكتملة / جزئية
		1	بيانات وصفية متاحة بالكامل	البيانات الوصفية الخاصة بمجموعات البيانات المستلمة مكتملة ويمكن الوصول إليها

(1) إمكانية الوصول (Accessibility)

(2) الدقة (Accuracy)

(3) التوصيف (Descriptiveness)

(4) الحداثة (Timeliness)

(5) الاكتمال (Completeness)

(6) التحقق من صحة البيانات (Validation)

(7) الشمولية (Coverage)

(8) الاتساق (Coherence)

تفاصيل معايير قياس مستوى جودة البيانات

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 الحداثة بتاريخ مرجعية البيانات

(1) إمكانية الوصول (Accessibility)

(2) الدقة (Accuracy)

(3) التوصيف (Descriptiveness)

(4) الحداثة (Timeliness)

(5) الاكتمال (Completeness)

(6) التحقق من صحة البيانات (Validation)

(7) الشمولية (Coverage)

(8) الاتساق (Coherence)

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
مستوى الحداثة	الحداثة هي ضمان توفير البيانات ضمن الفترة الزمنية المرجعية المحددة	0	بيانات غير حديثة	الفترة المرجعية للبيانات المستلمة أقدم من الفترة المرجعية المحددة للبيانات المطلوبة تتعدى العامين
		0.5	بيانات متوسطة الحداثة	الفترة المرجعية للبيانات المستلمة أقدم من الفترة المرجعية المحددة للبيانات المطلوبة من عام إلى عامين
		1	بيانات حديثة	البيانات المستلمة تمثل فترة مرجعية لا تتعدى عام واحد

تفاصيل معايير قياس مستوى جودة البيانات

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.

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
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	قيم مفقودة مؤثرة	وجود مجموعة كبيرة من القيم المفقودة تؤثر على دقة البيانات بشكل كبير (نسبة القيم المفقودة تتعدى 20٪)
		0.5	قيم مفقودة غير مؤثرة	وجود قيم بسيطة مفقودة وتؤثر على دقة البيانات بشكل بسيط (نسبة القيم المفقودة من 5٪ إلى 20٪)
		1	لا يوجد قيم مفقودة	لا توجد أية قيم مفقودة في البيانات (نسبة القيم المفقودة أقل 5٪)

(1) إمكانية الوصول (Accessibility)

(2) الدقة (Accuracy)

(3) التوصيف (Descriptiveness)

(4) الحداثة (Timeliness)

(5) الاكتمال (Completeness)

(6) التحقق من صحة البيانات (Validation)

(7) الشمولية (Coverage)

(8) الاتساق (Coherence)

تفاصيل معايير قياس مستوى جودة البيانات

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 المصدرية لدى الجهة الاتحادية.

(1) إمكانية الوصول (Accessibility)

(2) الدقة (Accuracy)

(3) التوصيف (Descriptiveness)

(4) الحداثة (Timeliness)

(5) الاكتمال (Completeness)

(6) التحقق من صحة البيانات (Validation)

(7) الشمولية (Coverage)

(8) الاتساق (Coherence)

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
التدقيق والتحقق من صحة البيانات	القيام بتطبيق قواعد التحقق من صحة البيانات (Data Validation Rules) على مستوى الأنظمة الإلكترونية وقواعد البيانات	0	لا يتم تطبيق قواعد التدقيق والتحقق	لا يوجد أي قواعد للتحقق من صحة البيانات مطبقة على مستوى الأنظمة أو قواعد البيانات لدى الجهة الاتحادية
		0.5	يتم تطبيق قواعد التدقيق والتحقق بشكل جزئي	يتم تطبيق بعض قواعد التدقيق والتحقق من صحة البيانات على مستوى الأنظمة أو قواعد البيانات لدى الجهة الاتحادية بشكل جزئي
		1	يتم تطبيق قواعد التدقيق والتحقق بشكل كامل	يتم تطبيق قواعد التدقيق والتحقق من صحة البيانات على مستوى الأنظمة أو قواعد البيانات بشكل كامل لدى الجهة الاتحادية

تفاصيل معايير قياس مستوى جودة البيانات

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:

(1) إمكانية الوصول (Accessibility)

(2) الدقة (Accuracy)

(3) التوصيف (Descriptiveness)

(4) الحداثة (Timeliness)

(5) الاكتمال (Completeness)

(6) التحقق من صحة البيانات (Validation)

(7) الشمولية (Coverage)

(8) الاتساق (Coherence)

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
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 غير مكتملة	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

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
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 غير مكتملة	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

تفاصيل معايير قياس مستوى جودة البيانات

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:

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
الاتساق الزمني	مقارنه التغيير المنطقي في البيانات عبر الزمن وعدم وجود تغيرات كبيرة أو غير منطقية أو غير مبررة	0	بيانات غير متسقة مؤثرة	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
		0.5	بيانات غير متسقة وغير مؤثرة	نسبة التأثير بسيطة بسبب عدم الاتساق في البيانات عند مقارنتها مع البيانات التاريخية السابقة، ولها تأثير بسيط على جودة البيانات لبعض الخصائص
		1	بيانات متسقة	لا يوجد حالات عدم اتساق في البيانات عند مقارنتها مع البيانات التاريخية السابقة لجميع الخصائص

(1) إمكانية الوصول (Accessibility)

(2) الدقة (Accuracy)

(3) التوصيف (Descriptiveness)

(4) الحداثة (Timeliness)

(5) الاكتمال (Completeness)

(6) التحقق من صحة البيانات (Validation)

(7) الشمولية (Coverage)

(8) الاتساق (Coherence)

تفاصيل معايير قياس مستوى جودة البيانات

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:

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
الاتساق مع مصادر أخرى	استخدام المصادر الرئيسية للبيانات من الجهة الرئيسية التي تجمع البيانات واتساقها مع قواعد البيانات المتوفرة مثل تاريخ الميلاد كمصدر بيان رئيسي يجب ان يكون من الجهة التي تصدر شهادة الميلاد	0	بيانات غير متسقة مؤثرة	يوجد حالات كثيرة من عدم الاتساق في البيانات عند مقارنتها مع مصادر أخرى، ولها تأثير كبير على جودة البيانات لبعض الخصائص
		0.5	بيانات غير متسقة وغير مؤثرة	يوجد حالات بسيطة من عدم الاتساق في البيانات عند مقارنتها مع مصادر أخرى، ولها تأثير بسيط على جودة البيانات لبعض الخصائص
		1	بيانات متسقة	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

(1) إمكانية الوصول (Accessibility)

(2) الدقة (Accuracy)

(3) التوصيف (Descriptiveness)

(4) الحداثة (Timeliness)

(5) الاكتمال (Completeness)

(6) التحقق من صحة البيانات (Validation)

(7) الشمولية (Coverage)

(8) الاتساق (Coherence)

تفاصيل معايير قياس مستوى جودة البيانات

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:

المقياس	المفهوم	التقييم		
		القيمة	الوصف	التوضيح
الاتساق الجغرافي	منطقية البيانات عند مقارنها بين الإمارات والمناطق الجغرافية في الدولة، بحيث تكون البيانات متناسقة مع التغيير الجغرافي	0	بيانات غير متسقة مؤثرة	يوجد حالات كثيرة من عدم الاتساق في البيانات عند مقارنتها جغرافياً ولها تأثير كبير على جودة البيانات لبعض الخصائص
		0.5	بيانات غير متسقة وغير مؤثرة	يوجد حالات بسيطة من عدم الاتساق في البيانات عند مقارنتها جغرافياً ولها تأثير بسيط على جودة البيانات لبعض الخصائص
		1	بيانات متسقة	لا يوجد حالات / حالات بسيط جدا من عدم اتساق في البيانات عند مقارنتها جغرافياً لجميع الخصائص

(1) إمكانية الوصول (Accessibility)

(2) الدقة (Accuracy)

(3) التوصيف (Descriptiveness)

(4) الحداثة (Timeliness)

(5) الاكتمال (Completeness)

(6) التحقق من صحة البيانات (Validation)

(7) الشمولية (Coverage)

(8) الاتساق (Coherence)

للتواصل والاستفسارات: FCSADMI@fcsc.gov.ae

مؤشر نضج البيانات

الدورة الثانية 2021