

معايير قياس مستوى جودة البيانات السجلية

تم تقسيم المؤشر إلى 3 محاور رئيسية لإدارة البيانات في الجهات:



ملاحظة: أوزان المعايير قابلة للتعديل وفقا لنتائج التقييم في كل دورة

تعريف البيانات السجلية

تعرف البيانات السجلية 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 البيانات

الوزن	المعايير	الوزن	المحاور الفرعية	الوزن	المحاور الرئيسية
5%	1.1.1 سياسة تنظيم وإدارة البيانات المؤسسية	15%	1.1 حوكمة الإطار التنظيمي والتشغيلي لإدارة البيانات	30%	1- الإطار التنظيمي والتشغيلي لإدارة البيانات
5%	1.1.2 الهيكلية والتنظيم في حوكمة البيانات				
5%	1.1.3 إتفاقيات إتاحة البيانات				
5%	1.2.1 حصر مصادر البيانات	15%	1.2 تأصيل مصادر البيانات المؤسسية	30%	2- تعزيز جودة وحماية البيانات
10%	1.2.2 إدارة وتوثيق البيانات الوصفية				
15%	2.1.1 منهجية إدارة جودة البيانات	15%	2.1 ضمان جودة البيانات	30%	2- تعزيز جودة وحماية البيانات
5%	2.2.1 إطار تحديد وتصنيف البيانات	15%	2.2 ضمان وحماية خصوصية البيانات		
10%	2.2.2 منهجية وعمليات تصنيف خصوصية البيانات				
15%	3.1.1 إتاحة البيانات الأولية	30%	3.1 تكامل البيانات وإتاحتها	40%	3- إتاحة البيانات وتدقيقها
5%	3.1.2 تفعيل الربط الإلكتروني				
10%	3.1.3 دورية مشاركة البيانات				
5%	3.2.1 حصر مجموعات البيانات المفتوحة	10%	3.2 البيانات المفتوحة	40%	3- إتاحة البيانات وتدقيقها
5%	3.2.2 إعداد ونشر البيانات المفتوحة				

تطبيق 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 والدورية (Timeliness)
5. الاكتمال (Completeness)
6. الشمولية (Coverage)
7. الاتساق (Consistency)

آلية تقييم جودة البيانات

- يتم تقييم جودة البيانات التي يتم تحميلها عبر قنوات الربط الإلكتروني فقط إلى المركز الاتحادي للتنافسية والإحصاء، وفي حال عدم تحميل البيانات بشكل إلكتروني عبر قنوات الربط الإلكترونية المعتمدة يتم تقييم مستوى الامتثال لهذا المحور على أنه "غير مكتمل".
- يتكون كل معيار من معايير جودة البيانات من مقياس واحد أو أكثر.
- تتم احتساب مقياس من مقاييس الجودة على حدة بآلية محددة.

تفاصيل معايير قياس مستوى جودة البيانات

ينبغي على الجهات الاتحادية إتاحة البيانات بشكل إلكتروني عن طريق الربط الإلكتروني وتسهيل الوصول للبيانات بصيغ ملائمة (Format) وقابلة للتبادل الإلكتروني

التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	توفر البيانات ضمن صيغ تنسيقية (Format) بشكل إلكتروني عن طريق الربط الإلكتروني ومتوافقة وملائمة يمكن من خلالها التعامل مع البيانات وقراءتها إلكترونياً دون الحاجة لأدوات أو برمجيات خاصة	إمكانية الوصول والتعامل مع البيانات
الصيغ التنسيقية للبيانات المستلمة لا تتوافق مع الصيغ التنسيقية القياسية المعتمدة وفقاً لمتطلبات المركز في تبادل البيانات الإلكتروني (إذا لم يتم توفير البيانات عن طريق الربط الإلكتروني تقيم جميع المعايير الأخرى بصفر)	بيانات بصيغ تنسيقية غير متوافقة مع الربط الإلكتروني	0		
الصيغ التنسيقية للبيانات المستلمة تتوافق بشكل تام مع الصيغ التنسيقية القياسية المعتمدة وفقاً لمتطلبات مؤشر نضج البيانات الحكومي في عملية تبادل البيانات الإلكتروني	بيانات بصيغ تنسيقية متوافقة بالكامل مع الربط الإلكتروني	1		

1. إمكانية الوصول (Accessibility)

2. الدقة (Accuracy)

3. التوصيف (Descriptiveness)

4. الحدثة والدورية (Timeliness)

5. الاكتمال (Completeness)

6. الشمولية (Coverage)

7. الاتساق (Consistency)

تفاصيل معايير قياس مستوى جودة 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:

التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	تكرار الرقم أو الاسم المعرف الموحد للحالة على سبيل المثال: • تكرار الخصائص مثل الرقم التعريفي، أو الشركة... إلخ • تكرار الرقم التعريفي أو اسم الشركة مع اختلاف باقي العناصر أو المتغيرات أو الحقول	التكرار
وجود تكرارات كثيرة في البيانات تؤثر على دقتها بشكل كبير (نسبة التكرار تتعدى 20%)	تكرار مؤثر	0		
وجود تكرارات بسيطة في البيانات تؤثر على دقتها بشكل بسيط (نسبة التكرار من 5% إلى 20%)	تكرار غير مؤثر	0.5		
لا يوجد أي تكرار في البيانات (نسبة التكرار أقل من 5%)	لا يوجد تكرار	1		

1. إمكانية الوصول (Accessibility)

2. الدقة (Accuracy)

3. التوصيف (Descriptiveness)

4. الحدثة والدورية (Timeliness)

5. الاكتمال (Completeness)

6. الشمولية (Coverage)

7. الاتساق (Consistency)

التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	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 الخانات الرقمية مثل العمر أو القيم التي تكون خارج المدى المتوقع أو المنطق	القيم الخطأ
وجود مجموعة كبيرة من القيم الخاطئة والتي تؤثر على دقة البيانات بشكل كبير (نسبة القيم الخاطئة تتعدى 20%)	قيم خاطئة مؤثرة	0		
وجود مجموعة بسيطة من القيم الخاطئة والتي تؤثر على دقة البيانات بشكل بسيط (نسبة القيم الخاطئة من 5% إلى 20%)	قيم خاطئة غير مؤثرة	0.5		
لا توجد أية قيم خاطئة في البيانات (نسبة القيم الخاطئة أقل من 5%)	لا يوجد قيم خاطئة	1		

تفاصيل معايير قياس مستوى جودة 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:

التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	تطبيق التصنيفات الوطنية والدولية المعتمدة لتجنب وجود اختلافات في التصنيف المعتمدة من ناحية الترميز والوصف على سبيل المثال: • ISIC-2، ISIC-4، الخ، وفق دليل المعايير الوطنية للبيانات الإحصائية	1. إمكانية الوصول (Accessibility) 2. الدقة (Accuracy) 3. التوصيف (Descriptiveness) 4. الحدثة والدورية (Timeliness) 5. الاكتمال (Completeness) 6. الشمولية (Coverage) 7. الاتساق (Consistency)
غير متبع للتصنيفات الوطنية أو الدولية المعتمدة	غير متبع	0		
متبع بشكل جزئي أو نسخة قديمة لتصنيف يؤثر على البيانات	متبع بشكل جزئي	0.5		
متبع لنسخة حديثة أو مقبولة للتصنيفات الوطنية أو الدولية	متبع بشكل كامل	1		
التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	القيام بتطبيق قواعد التحقق من صحة البيانات (Data Validation Rules) على مستوى الأنظمة الإلكترونية وقواعد البيانات (هي عبارة عن فحص منطقية العلاقة بين أكثر من متغير مثلاً العمر مع المستوى التعليمي)	5. الاكتمال (Completeness) 6. الشمولية (Coverage) 7. الاتساق (Consistency)
لا يوجد أي قواعد للتحقق من صحة البيانات مطبقة على مستوى الأنظمة أو قواعد البيانات لدى الجهة الاتحادية	لا يتم تطبيق قواعد التدقيق والتحقق	0		
يتم تطبيق بعض قواعد التدقيق والتحقق من صحة البيانات على مستوى الأنظمة أو قواعد البيانات لدى الجهة الاتحادية بشكل جزئي	يتم تطبيق قواعد التدقيق والتحقق بشكل جزئي	0.5		
يتم تطبيق قواعد التدقيق والتحقق من صحة البيانات على مستوى الأنظمة أو قواعد البيانات بشكل كامل لدى الجهة الاتحادية	يتم تطبيق قواعد التدقيق والتحقق بشكل كامل	1		

تفاصيل معايير قياس مستوى جودة البيانات

ينبغي على الجهات الاتحادية توثيق وإتاحة قابلية الاطلاع على البيانات الوصفية الشاملة لمجموعات البيانات التي يتم مشاركتها وفقاً للمعيار التالي:

التقييم			المفهوم	المقياس	
التوضيح	الوصف	القيمة	إتاحة البيانات الوصفية (Metadata) من قبل الجهة المزودة للبيانات وإرفاق جميع الأدلة والتصانيف التي استخدمت في تجميع البيانات وذلك لضمان تسهيل الوصول والاطلاع عليها من قبل المستخدمين من مجموعات البيانات	إمكانية الاطلاع على البيانات الوصفية	
البيانات الوصفية الخاصة بمجموعات البيانات المستلمة غير متاحة	بيانات وصفية غير متاحة	0			
البيانات الوصفية الخاصة بمجموعات البيانات المستلمة غير مكتملة / جزئية	بيانات وصفية متاحة بشكل جزئي	0.5			
البيانات الوصفية الخاصة بمجموعات البيانات المستلمة مكتملة ويمكن الوصول إليها	بيانات وصفية متاحة بالكامل	1			

1. إمكانية الوصول (Accessibility)

2. الدقة (Accuracy)

3. التوصيف (Descriptiveness)

4. الحدثة والدورية (Timeliness)

5. الاكتمال (Completeness)

6. الشمولية (Coverage)

7. الاتساق (Consistency)

تفاصيل معايير قياس مستوى جودة 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

التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	ضمان توفير بيانات حديثة مطابقة للفترة الزمنية المرجعية المحددة مثال على ذلك: • بيانات المشتغلين 2020 والتي تمثل الفترة الزمنية من 2020-1-1 إلى 2020-12-31 ويتم تسليمها في 2021	الحداثة
الفترة المرجعية للبيانات المستلمة أقدم من الفترة المرجعية المحددة للبيانات المطلوبة تتعدى العام	قديم	0		
الفترة المرجعية للبيانات المستلمة (عام واحد أو أقل) عن الفترة المرجعية المحددة للبيانات المطلوبة	بيانات سنوية	0.5		
الفترة المرجعية للبيانات المستلمة (ربع سنوي أو أقل)	ربع سنوي	1		
التقييم			المفهوم	المقياس
معادلة القياس			ضمان توفر البيانات بشكل آني ومستمر يضمن دورية البيانات وتدققها خلال الفترة الزمنية المرجعية المحددة	الدورية
$\frac{\text{دورية البيانات المستلمة}}{\text{دورية البيانات المحددة في الاتفاقية}} =$ <p>مثال على ذلك: في حال ان كانت الدورية المعتمدة نصف سنويه وتم استلام البيانات مرة واحدة فتكون النتيجة ½</p>				

1. إمكانية الوصول (Accessibility)

2. الدقة (Accuracy)

3. التوصيف (Descriptiveness)

4. الحداثة والدورية (Timeliness)

5. الاكتمال (Completeness)

6. الشمولية (Coverage)

7. الاتساق (Consistency)

تفاصيل معايير قياس مستوى جودة 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

التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	توفير البيانات التاريخية لضمان توفر سلاسل زمنية وتاريخية مكتملة	البيانات التاريخية
توفير بيانات تاريخية (سنة واحدة سابقة)	لا توجد سلسلة زمنية	25		
توفير بيانات تاريخية (3 سنوات سابقة)	سلسلة زمنية صغيرة	50		
توفير بيانات تاريخية (5 سنوات سابقة)	سلسلة زمنية متوسطة	75		
توفير بيانات تاريخية (10 سنوات سابقة)	سلسلة زمنية مكتملة	100		

1. إمكانية الوصول (Accessibility)

2. الدقة (Accuracy)

3. التوصيف (Descriptiveness)

4. الحداثة والدورية (Timeliness)

5. الاكتمال (Completeness)

6. الشمولية (Coverage)

7. الاتساق (Consistency)

تفاصيل معايير قياس مستوى جودة البيانات

ينبغي أن تكون البيانات التي يتم مشاركتها متكاملة ويمكن تفسيرها بشكل منطقي وواضح، وأن تكون قابلة للاستخدام دون الحاجة إلى بيانات مرجعية إضافية لتفسيرها أو استخدامها.

التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	الخانات أو الحقول في مجموعة البيانات التي لا تحتوي على أي قيمة، أي أن بعض العناصر أو المتغيرات لم يتم استكمال تعبئتها	القيم المفقودة
وجود مجموعة كبيرة من القيم المفقودة تؤثر على دقة البيانات بشكل كبير (نسبة القيم المفقودة تتعدى 20%)	قيم مفقودة مؤثرة	0		
وجود قيم بسيطة مفقودة وتؤثر على دقة البيانات بشكل بسيط (نسبة القيم المفقودة من 5% إلى 20%)	قيم مفقودة غير مؤثرة	0.5		
لا توجد أية قيم مفقودة في البيانات (نسبة القيم المفقودة أقل من 5%)	لا يوجد قيم مفقودة	1		

1. إمكانية الوصول (Accessibility)

2. الدقة (Accuracy)

3. التوصيف (Descriptiveness)

4. الحدثة والدورية (Timeliness)

5. الاكتمال (Completeness)

6. الشمولية (Coverage)

7. الاتساق (Consistency)

تفاصيل معايير قياس مستوى جودة 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:

التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	تغطية البيانات لكافة المناطق الجغرافية (الجيومكانية)	(الجيومكانية) على مستوى دولة الإمارات
توجد إمارة أو أكثر لم يتم تغطيتها ضمن نطاق البيانات	تغطية ضعيفة	0		
تم تغطية جميع الإمارات ضمن نطاق البيانات مع وجود نقص في بيانات بعض المناطق الرئيسية والفرعية	تغطية غير مكتملة	0.5		
تم تغطية جميع الإمارات والمناطق الرئيسية والفرعية المطلوبة بشكل كامل	تغطية كاملة	1		

1. إمكانية الوصول (Accessibility)

2. الدقة (Accuracy)

3. التوصيف (Descriptiveness)

4. الحدثة والدورية (Timeliness)

5. الاكتمال (Completeness)

6. الشمولية (Coverage)

7. الاتساق (Consistency)

التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	تغطية البيانات لجميع الفئات والمتغيرات المستهدفة، على سبيل المثال: • فئة المناطق الحرة والجهات الشبه حكومية • الفئات الاقتصادية كالأنشطة الاقتصادية وعدد العاملين وغيرها. • الخصائص السكانية كالنوع والحالة الاجتماعية وغيرها	التغطية والشمولية
توجد فئة رئيسية أو أكثر لم يتم تغطيتها ضمن نطاق البيانات المطلوبة	تغطية ضعيفة	0		
تم تغطية جميع الفئات مع وجود نقص بسيط في بعضها وتأثيرها بسيط على شمولية البيانات	تغطية غير مكتملة	0.5		
تم تغطية جميع الفئات في البيانات بشكل كامل	تغطية كاملة	1		

تفاصيل معايير قياس مستوى جودة 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:

التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	مقارنه التغير المنطقي في البيانات عبر الزمن وعدم وجود تغيرات كبيرة، أو غير منطقية، أو غير مبررة أو مختلفة مثال: • مقارنة مؤشر أسعار الإيجارات في 2022 مع مؤشر أسعار ايجارات 2023	الاتساق الزمني
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	بيانات غير متسقة مؤثرة	0		
نسبة التأثير بسيطة بسبب عدم الاتساق في البيانات عند مقارنتها مع البيانات التاريخية السابقة، ولها تأثير بسيط على جودة البيانات لبعض الخصائص	بيانات غير متسقة وغير مؤثرة	0.5		
لا يوجد حالات عدم اتساق في البيانات عند مقارنتها مع البيانات التاريخية السابقة لجميع الخصائص	بيانات متسقة	1		

1. إمكانية الوصول (Accessibility)

2. الدقة (Accuracy)

3. التوصيف (Descriptiveness)

4. الحدائة والدورية (Timeliness)

5. الاكتمال (Completeness)

6. الشمولية (Coverage)

7. الاتساق (Coherence)

تفاصيل معايير قياس مستوى جودة 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:

التقييم			المفهوم	المقياس
التوضيح	الوصف	القيمة	يجب التأكد من ان البيانات منطقية من خلال مقارنتها بمصادر أخرى رسمية بحيث لا يوجد فوارق كبيرة مثال: • مقارنة بيانات سجل العاملين مع بيانات سجل مسح القوى العاملة	الاتساق مع مصادر أخرى
يوجد حالات كثيرة من عدم الاتساق في البيانات عند مقارنتها مع مصادر أخرى، ولها تأثير كبير على جودة البيانات لبعض الخصائص	بيانات غير متسقة مؤثرة	0		
يوجد حالات بسيطة من عدم الاتساق في البيانات عند مقارنتها مع مصادر أخرى، ولها تأثير بسيط على جودة البيانات لبعض الخصائص	بيانات غير متسقة وغير مؤثرة	0.5		
لا يوجد حالات عدم اتساق في البيانات عند مقارنتها مع مصادر أخرى لجميع المتغيرات	بيانات متسقة	1		

1. إمكانية الوصول (Accessibility)

2. الدقة (Accuracy)

3. التوصيف (Descriptiveness)

4. الحدثة والدورية (Timeliness)

5. الاكتمال (Completeness)

6. الشمولية (Coverage)

7. الاتساق (Coherence)

تفاصيل معايير قياس مستوى جودة 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:

التقييم			المفهوم	المقياس		
التوضيح	الوصف	القيمة	الاتساق الجغرافي			
يوجد حالات كثيرة من عدم الاتساق في البيانات عند مقارنتها جغرافياً ولها تأثير كبير على جودة البيانات لبعض الخصائص	بيانات غير متسقة مؤثرة	0			يجب التأكد من منطقية البيانات عند مقارنتها بين المناطق الجغرافية في الدولة أو مع دول أخرى، بحيث يتم التأكد من أن البيانات متناسقة مع المتغير الجغرافي	
يوجد حالات بسيطة من عدم الاتساق في البيانات عند مقارنتها جغرافياً ولها تأثير بسيط على جودة البيانات لبعض الخصائص	بيانات غير متسقة وغير مؤثرة	0.5				
لا يوجد حالات / حالات بسيط جداً من عدم اتساق في البيانات عند مقارنتها جغرافياً لجميع الخصائص	بيانات متسقة	1				

1. إمكانية الوصول (Accessibility)

2. الدقة (Accuracy)

3. التوصيف (Descriptiveness)

4. الحدثة والدورية (Timeliness)

5. الاكتمال (Completeness)

6. الشمولية (Coverage)

7. الاتساق (Coherence)



   @FCSCUAE  @FCSAUAE