



دليل الممارس من insyt2

التقييمات السيكومترية

كل ما يحتاج متخصصو الموارد البشرية إلى معرفته

دليل منهجي وعملي، صُمم لبيئة الخليج، ويغطي علم القياس، وتصميم التقييم، والعدالة، والحوكمة، والتكنولوجيا والممارسة المهنية.



يجب ألا ينتهي التقييم بدرجة.

بل ينبغي أن يعزز قراراً، ويحسن حواراً، أو يخلق خطوة تالية واضحة في التطوير.

01

السياق

الدور، القرار، والفئة المستهدفة

02

الدليل

أساليب شفافة ومناسبة

03

الفعل

خطوة تالية واضحة ومرئية

ما الذي يتضمنه هذا الدليل؟

مسار منظم يبدأ بالمبادئ الأساسية وينتهي بالممارسة المهنية المتقدمة.

01

الأسس | الفصول 01-04

ماهية التقييم السيكمومي؛ طبيعة المهنة؛ أساليب التقييم؛ المخاطر.

02

علم القياس | الفصول 05-08

البُنى؛ الثبات؛ الصدق؛ المعايير؛ العدالة، وحدود القرار.

03

دورة حياة التقييم | الفصول 09-14

الاستراتيجية؛ الموزدون؛ التطوير؛ الإدارة؛ التسجيل؛ دمج الأدلة.

04

أساليب التقييم | الفصول 15-17

الأداء الأقصى؛ الأداء المعتاد؛ أدوات التطوير والتنظيم.

05

الحوكمة والمستقبل الرقمي | الفصول 18-20

الأخلاقيات؛ القانون؛ الذكاء الاصطناعي؛ مركز التميز.

06

بناء المسيرة المهنية | الفصول 21-25

المسارات؛ الكفاءة؛ المؤهلات؛ خارطة الطريق؛ المحفظة.

معيار insyt2: نبدأ بالسياق، نطلب الدليل، ونحوّل النتيجة إلى فعل مهني مسؤول.

ممارسة متخصصة في التقييم والتطوير، تركيزها الأساسي على منطقة الخليج.



تأسست insyt2 في عام 2020، وتربط بين علم النفس المهني، واستشارات الموارد البشرية، وتطوير القيادات، والتدريب التنفيذي - من الفرد إلى المؤسسة بأكملها.

لماذا تعمل المؤسسات معنا؟

01

مصمم حسب الحاجة

نبني الحل حول القرار، والدور، والسياق.

02

فهم خليجي

ملاءمة إقليمية مدعومة بانضباط سيكومتري عالمي.

03

ثنائي اللغة

تصمم الإنجليزية والعربية معاً منذ البداية.

04

متكامل

نربط التقييم بالتقارير والتطوير.

05

قائم على الدليل

أساليب شفافة وحكم مهني مسؤول.

06

جاهز للفعل

كل مخرج يجب أن يقود إلى خطوة عملية تالية.

في insyt2 نؤمن بأن التقييم ينبغي أن يقوّي القرارات والحوارات والتطوير - لا أن يكتفي بإنتاج درجة.

أسئلة أفضل. أدلة أفضل. حكم مهني أفضل.



المشكلة التي نريد معالجتها

توجد منتجات كثيرة في السوق توصف بأنها «سيكومترية» أو «علمية» أو «متحقق من صدقها» من دون وجود أدلة كافية. وقد تخلق الملفات الجذابة ولوحات المعلومات ثقة لا تستند إلى وضوح البنية، أو الثبات، أو الصدق، أو العدالة، أو أدلة محلية مناسبة.

ما الذي نريد من هذا الدليل أن يحققه؟

نريد مساعدة متخصصي الموارد البشرية على طرح أسئلة أفضل، والتمييز بين الدليل والتسويق، والتعاقد على التقييمات وتفسيرها بمسؤولية، واستخدام التقييمات والاستبيانات بطرق عادلة، ومتناسبة، وذات مصداقية ثقافية، ومفيدة فعلاً.

معيار الممارسة لدى insyt2

01

السياق

ابدأ بالدور، والقرار، والفئة المستهدفة.

02

الدليل

استخدم أساليب شفافة ومصادر متعددة.

03

الفعل

اجعل الخطوة التالية واضحة ومرئية.

04

التطوير

استخدم التقييم لتوسيع القدرات، لا لتثبيت الهوية.

في insyt2 نؤمن بأن... يجب أن يتناسب معيار الدليل مع القرار، والفئة المستهدفة، وعواقب الخطأ. **ولن نقوم أبداً ب...** اعتبار لوحة معلومات مصقولة أو تسمية جذابة أو توصية بالذكاء الاصطناعي دليلاً على الجودة السيكومترية.

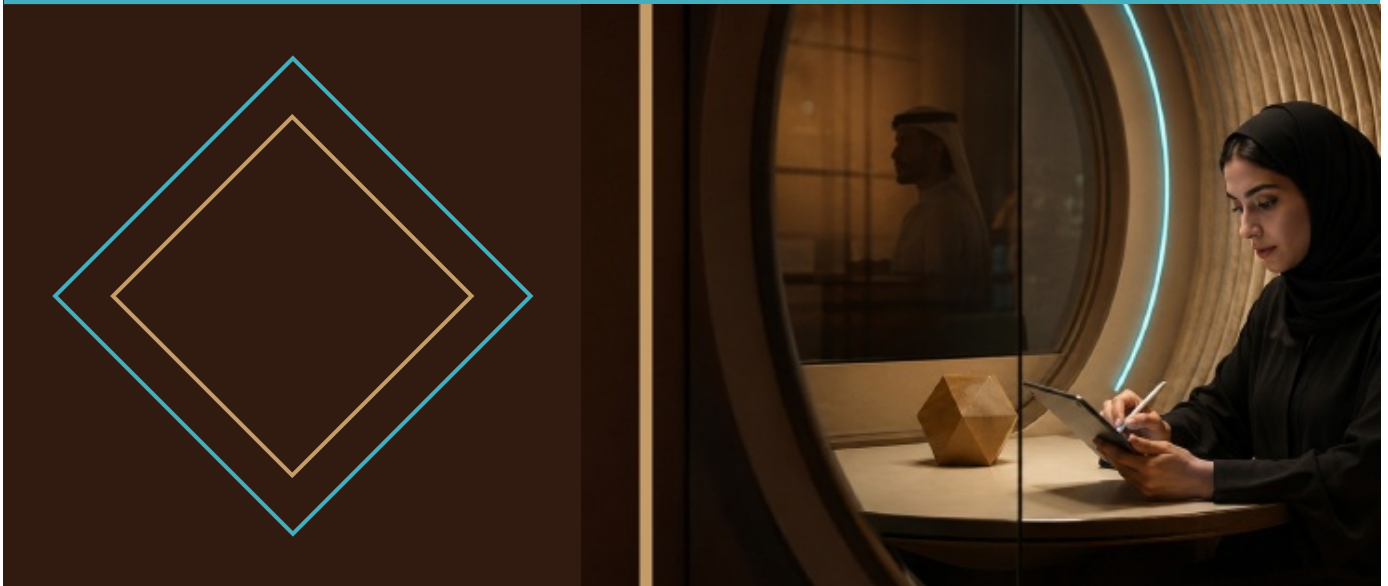
قائمة عملية لضبط الجودة عند شراء التقييمات أو تصميمها أو تفسيرها أو حوكمتها.

01 ابدأ بالقرار، لا بالأداة. عرّف القرار، والفئة المستهدفة، وفجوة الأدلة قبل اختيار أي اختبار.	02 الدرجة تقدير وليست حقيقة مطلقة. كل درجة تتضمن قدراً من عدم اليقين؛ والفروق الصغيرة نادراً ما تبرز ترتيباً دقيقاً بين الأفراد.	03 الصدق يتعلق بالتفسير والاستخدام. يجب أن تدعم الأدلة المعنى والنتائج المترتبة على الاستخدام في السياق المحدد.
04 الثبات ضروري لكنه غير كافٍ. قد يكون القياس متسقاً، ومع ذلك يكون غير ذي صلة، أو منحازاً، أو سيئ التفسير.	05 تحليل الوظيفة هو أساس الاختيار. اربط متطلبات العمل بالبُنى والمحتوى والقرارات بصورة صريحة.	06 العدالة تُصمم عبر دورة الحياة كاملة. إمكانية الوصول، والمحتوى، والتسجيل، والتنبؤ، والقرارات، والمراقبة كلها مهمة.
07 يمكن أن يتعايش التوحيد والتيسيرات. احم إمكانية الوصول إلى البنية المقصودة، ولا تفترض أن العدالة تعني المعاملة المتطابقة.	08 قد تكون المعايير مضللة. يجب أن تكون المجموعات المرجعية مناسبة، وحديثة، وموصوفة بشفافية.	09 حدود القرار تخلق عواقب. تحتاج العتبات إلى أدلة، وتحليل للأخطاء، وحوكمة واضحة.
10 المزيد من التقييم ليس أفضل تلقائياً. أضف أساليب جديدة فقط عندما تقدم معلومات مختلفة وذات صلة بالعمل.	11 العلامة التجارية للمورد ليست دليلاً فنياً. اطلب الأدلة الفنية، والدراسات، ونتائج المجموعات الفرعية، والضوابط، وسجل الإصدارات.	12 الدليل المحلي مهم. الأدلة المنشورة نقطة بداية؛ راقب أداء الأداة في سياقك أنت.
13 الذكاء الاصطناعي لا يُعفى من المعايير السيكمترية. النماذج الجديدة تحتاج أيضاً إلى صلة بالعمل، وثبات، وصدق، وعدالة.	14 الكفاءة تشمل معرفة حدودك. اطلب الإشراف أو الدعم المتخصص عندما تتجاوز المهمة نطاق خبرتك.	15 تجربة المرشح جزء من الجودة. الوضوح، والكرامة، وإمكانية الوصول، والدعم، والمراجعة كلها تؤثر في قابلية الدفاع عن العملية.

في insyt2 نؤمن بأن أقوى الممارسين يستطيعون التحدث بلغة الإحصاء ولغة الأعمال معاً. الدليل دون سياق ضوضاء؛ والسياق دون دليل رأي؛ والحكم المهني يصل بينهما.

الأسس

ماهية المجال، ولماذا هو مهم، وأين يقع متخصص الموارد البشرية داخله.



في هذا الجزء

01 ما هو التقييم السيكمي - وما الذي لا يُعد كذلك

02 المشهد المهني

03 أساليب التقييم وحالات الاستخدام

04 مستوى المخاطر والتناسب

العلم • العدالة • جودة القرار • الحكم المهني

رؤية insyt2

السؤال ليس: أي اختبار يبدو أكثر إبهاراً؟ بل: أي دليل يخدم القرار، والأشخاص المتأثرين به، وما سيحدث بعد ذلك بصورة أفضل؟

الفصل

01

ما هو التقييم السيكمتري - وما الذي لا يُعد كذلك

عملية قياس منضبطة، لا مجموعة من الاستبيانات الجذابة.

تعريف عملي

في insyt2 نؤمن بأن...

لن تقوم insyt2 أبداً بـ...

وصف أداة بأنها سيكمترية لمجرد أنها تنتج ملفاً شخصياً أو درجة أو لوحة معلومات جذابة.

ممارسة insyt2: السياق قبل المحتوى

لا نبدأ بالسؤال: أي اختبار ينبغي أن نشترى؟ نبدأ بالدور، والقرار، والفئة المستهدفة، وفجوة الأدلة، وعواقب الخطأ؛ ثم نختار الأسلوب أو نصممه.

الاختبار، والتقييم، والقرار، والحوكمة

المستوى	السؤال الأساسي	المخرجات المعتادة
الاختبار أو الأداة	كيف سيتم أخذ عينة من السمة واحتساب درجتها؟	مفردات، تمارين، مقاييس تقدير، درجات خام، ودرجات معيارية.
عملية التقييم	كيف سثجمع الأدلة بعدالة واتساق؟	تعليمات، بروتوكول إدارة، تيسيرات، تقارير، وفحوص جودة.
نظام القرار	كيف ستؤثر الأدلة في إجراء تنظيمي؟	حدود قرار، قواعد مركبة، أحكام اللجان، الاستثناءات، ومسار تدقيق.
نظام الحوكمة	من المسؤول عن الجودة والمخاطر والمراجعة؟	سياسة، أدوار، أدوات معتمدة، مراقبة، حوادث، ضبط التغيير، وإيقاف الأدوات.

ما الذي تستطيع التقييمات السيكمترية القيام به؟

- تقديم أدلة موحدة عن خصائص ذات صلة بالعمل يصعب ملاحظتها باتساق في حوار غير رسمي.
- تحسين قابلية المقارنة من خلال منح المرشحين أو الموظفين فرصاً متكافئة لإظهار القدرات أو الأنماط ذات الصلة.
- دعم التنبؤ، والتصنيف، وتشخيص الاحتياجات التطويرية، والتأمل المنظم عندما يكون الاستخدام المقصود مدعوماً بالأدلة.
- جعل قواعد القرار والافتراضات مرئية بما يسمح بالتدقيق والتقييم والتحسين.
- تقليل بعض أشكال عدم الاتساق البشري، مع إدخال مخاطر أخرى يجب ضبطها.

ما الذي لا تستطيع التقييمات السيكمترية القيام به؟

- الكشف عن حقيقة كاملة ودائمة ومستقلة عن السياق حول الشخص.
- إزالة الحكم البشري أو المساءلة أو عدم اليقين أو التحيز من قرار وظيفي.
- التنبؤ بالسلوك بصورة مثالية عبر جميع الأدوار والثقافات والفترات الزمنية والظروف التنظيمية.
- تعويض غموض تعريف الدور، أو ضعف معايير الأداء، أو خلل عملية اتخاذ القرار.
- تبرير تشخيص سريري أو استنتاج طبي عندما تكون الأداة مصممة للاستخدام المهني.
- تحويل قرار مشكوك فيه أخلاقياً إلى قرار مقبول لمجرد أن خوارزمية أنتجته.

القياس بدلاً من التصنيف

قد يقول تقرير مسؤول: «في ظروف اختبار الاستدلال العددي المحدد بزمان، كان أداء المرشح أعلى من متوسط مجموعة المقارنة، مع نطاق ثقة يقع تقريباً ضمن المستوى فوق المتوسط». أما التقرير غير المسؤول فيحوّل الدرجة إلى هوية: «هذا الشخص شديد الذكاء وسينجح». العبارة الأولى مقيدة بالبنية، والأسلوب، والظروف، والمجموعة المعيارية، وعدم اليقين. أما الثانية فتعمم أكثر مما تسمح به البيانات وتوحي بدرجة من اليقين لا يدعمها الدليل.

لن تخط insyt2 أبداً بين جودة العرض وجودة الدليل

لوحة معلومات أنيقة، أو اسم نمط يسهل تذكره، أو درجة تبدو شديدة الدقة لا تثبت الجودة السيكمترية. نبحث عن وضوح البنية، والثبات، والصدق، والعدالة، ودليل ملائم لحالة الاستخدام.

انضباط الأسئلة الخمسة

1. ما الذي يتم قياسه تحديداً؟
 2. إلى أي مدى يتم قياسه باتساق ودقة؟
 3. ما الأدلة التي تدعم التفسير والاستخدام المقصودين؟
 4. لمن، وتحت أي ظروف، وبالمقارنة مع أي مجموعة مرجعية؟
 5. ماذا يحدث عندما تكون الدرجة خاطئة، أو ناقصة، أو غير متاحة، أو أسوء استخداماً؟
- هذه الأسئلة هي بداية الثقافة السيكمترية، كما أنها حماية عملية من شراء الأدوات بسبب حداثتها، أو شهادات العملاء، أو عرض بيعي مقنع.

الفصل

02

المشهد المهني

يؤدي مستخدمو الاختبارات، وأخصائيو التقييم، والسيكمتريون، وعلماء النفس، والمحللون، وقادة المنتجات، ومختصو الحوكمة أعمالاً مختلفة.

مجال واحد، ومستويات متعددة من الممارسة

في insyt2 نؤمن بأن...

لن نقوم insyt2 أبداً بـ..

افتراض أن الإلمام بأداة واحدة يخلق تلقائياً كفاءة في كامل تخصص القياس السيكمتري.

يتعامل متخصصو الموارد البشرية غالباً مع السيكمترية لأول مرة كمشرفين على الإدارة، أو مستلمين للتقارير، أو مشترين للأدوات. ويمكن للمسيرة المهنية بعد ذلك أن تتجه في عدة مسارات. الفارق الأساسي ليس المكانة الوظيفية، بل طبيعة القرارات التي تملك الكفاءة والصلاحيات لاتخاذها. فإدارة أداة منشورة وفق دليلها تتطلب خبرة مختلفة عن بناء نموذج تسجيل، أو التحقق من نظام اختيار، أو تصميم بنك مفردات.

الدور	المساهمة المعتادة	العمق الفني	الحد الذي يجب احترامه
منسق التقييم	الجدولة، والتواصل مع المشاركين، ودعم المنصة، والسجلات، والتعامل مع الحوادث.	تأسيسي	لا يفسر الدرجات المعقدة بصورة مستقلة ولا يغير الشروط المعيارية.
مستخدم اختبارات مؤهل	يختار الأدوات المنشورة المناسبة، ويديرها، ويسجلها، ويفسرها، ويقدم التغذية الراجعة.	تطبيقي	يعمل ضمن مؤهلاته، وقواعد الناشر، ونطاق الممارسة المحلي.
أخصائي تقييم المواهب	يصمم عمليات متعددة الأساليب، ويدرب المقيمين، وينصح الموارد البشرية، ويراقب النتائج.	تطبيقي إلى متقدم	يصعد القضايا المعقدة في التحقق أو القانون أو الممارسة السريرية أو الإحصاء.
استشاري أو مصمم تقييم	يبنى ملفات الوظائف، والتمارين، واختبارات الحكم الموقفي، والاستبيانات، وأدلة التسجيل.	متقدم	يوثق الأدلة ويتجنب الادعاءات المملوكة التي لا تسندها أدلة.
سيكمتري	يطور المقاييس ويقيم الثبات والصدق والتكافؤ والربط ونماذج المفردات والتسجيل.	تخصص كمي	يحتاج إلى خبرة عميقة في القياس والإحصاء.
أخصائي علم نفس صناعي-تنظيمي أو مهني	يطبق علم النفس في الاختيار والتطوير والفرق والعمل والمؤسسات.	مهنة واسعة	تختلف الألقاب والتسجيل والأنشطة المسموح بها حسب الولاية القضائية.
مدير منتج تقييم	يربط احتياجات المستخدم بالعلم وتجربة الاستخدام والتكنولوجيا والخصوصية والأولويات التجارية.	عابر للتخصصات	لا يجوز التضحية بالضوابط العلمية أو القانونية من أجل السرعة.
مسؤول حوكمة التقييم أو ضمان الذكاء الاصطناعي	يضع السياسة، وفئات المخاطر، والموافقات، والتدقيق، والمراقبة، والمساءلة.	مخاطر عابرة للتخصصات	يحتاج إلى مراجعة مستقلة وإمكانية الوصول إلى الأدلة الفنية.

الكفاءة مرتبطة بالمهمة

قد يكون الشخص ممتازاً في تصميم المقابلات المنظمة ولا يملك خبرة كبيرة في نظرية الاستجابة للمفردة. وقد يكون آخر سيكمترياً قوياً لكنه أقل خبرة في تقديم التغذية الراجعة أو تنفيذ الحلول داخل المؤسسات. لذلك تتطلب الممارسة الأخلاقية خريطة للكفاءات، لا تسمية واحدة. قبل قبول أي مهمة، اسأل ما إذا كنت تملك المعرفة والخبرة الخاضعة للإشراف والأدوات والصلاحيات اللازمة لكل عنصر من عناصر المشروع.

المهمة	أسئلة الحد الأدنى للكفاءة
اختيار اختبار منشور	هل أستطيع تعريف البنية والاستخدام، وقراءة الدليل الفني، وتقييم المعايير والصدق والثبات والعدالة والملاءمة العملية؟
تفسير ملف فردي	هل أفهم معنى الدرجات، وعدم اليقين، والتفاعلات بين المقاييس، والقيود، والتغذية الراجعة المناسبة؟
تحديد حد قرار للاختيار	هل أستطيع ربط العتبة بمتطلبات الوظيفة، وتكاليف الأخطاء، والسياق القانوني، وأدلة التحقق؟
بناء استبيان	هل أستطيع تعريف المجال، وكتابة المفردات ومراجعتها، وتجريبها، وتحليل بنيتها وثباتها، والتحقق من التفسيرات؟
بناء درجة خوارزمية	هل يستطيع الفريق حوكمة مصدر البيانات، وتسرب المعلومات، وفرط الملاءمة، وأداء المجموعات الفرعية، وقابلية التفسير، والانحراف بمرور الوقت، والأمن، والإشراف البشري؟
ترجمة اختبار أو تكييفه	هل يستطيع الفريق معالجة التكافؤ اللغوي والثقافي والبنائي والإداري والقياسي، بدلاً من مجرد ترجمة الكلمات؟

الألقاب المهنية والمؤهلات

تختلف الألقاب المهنية، وأنظمة التسجيل، ومؤهلات استخدام الاختبارات بين الدول. قد تمنح شهادة من ناشر ما صلاحية استخدام أداة بعينها، لكنها لا تثبت تلقائياً كفاءة واسعة في السيكمترية. وبالمقابل، قد يوفر مؤهل دراسات عليا أساساً نظرياً قوياً من دون أن يمنح صلاحية استخدام أدوات مقيدة أو يثبت وجود ممارسة حديثة تحت الإشراف. ابن محفظة أدلة مترابطة تشمل التعليم، والمؤهلات المعترف بها، وتدريب الناشرين، والحالات تحت الإشراف، والعمل الفني، والتطوير المهني المستمر، والمساءلة الأخلاقية.

منظور insyt2 للمسيرة المهنية: اختر العمل أولاً، ثم المؤهل

حدد نوع العمل في التقييم الذي تريد أن تصبح موثقاً في أدائه. فمسار التعلم لمستخدم اختبارات مؤهل يختلف عن مسار السيكمترية أو مسؤول الحوكمة أو أخصائي المنتج.

حدود النطاق التي تحمي الأفراد والمسارات المهنية

- لا تستخدم أدوات مهنية لتشخيص اضطرابات نفسية أو حالات صحية.
- لا تستنتج خصائص محمية أو حساسة أو طبية إلا عندما يكون الاستخدام قانونياً وضرورياً وصريحاً ومبرراً مهنياً.
- لا تفسر أدوات مقيدة دون المؤهل أو الإشراف المطلوب.
- لا تخلق مظهراً زائفاً للدقة من خلال الإبلاغ عن درجات أو معايير أو خوارزميات لا تدعمها أدلة.
- لا تعد بأن التقييم «خالي من التحيز» أو «موضوعي بالكامل». صف الأدلة والضوابط والمخاطر المتبقية والمراقبة.
- لا تسمح للملكية التجارية للأداة بأن تمنع التقييم الفني الشفاف.

الفصل

03

أساليب التقييم وحالات الاستخدام

تأخذ الأساليب المختلفة عينات من أشكال مختلفة من السلوك، وتجيب عن أسئلة تنظيمية مختلفة.

تميزين أساسيين

لن تقوم insyt2 أبداً بـ...

إعادة استخدام أداة واحدة في التوظيف والترقية والتطوير لمجرد أن محتواها يبدو ذا صلة.

في insyt2 نؤمن بأن...

تقيس أدوات الأداء الأقصى ما يستطيع الشخص القيام به تحت شروط محددة. وتشمل اختبارات الاستدلال، والمعرفة، والمهارات، وعينات العمل، والمحاكاة، والعديد من اختبارات الحكم الموقفي. أما أدوات الأداء المعتاد فتسأل كيف يميل الشخص إلى التصرف أو التفضيل أو التقدير أو الاستجابة، وتشمل الشخصية، والدوافع، والاهتمامات، والقيم، والاتجاهات. تختلف متطلبات الاستعداد، واحتمالات تحريف الاستجابة، والوقت، والتسجيل، والثبات، والصدق، والتغذية الراجعة بين هاتين العائلتين.

والتمييز الثاني هو بين التقرير الذاتي والأداء المظهر فعلياً. يتيح التقرير الذاتي الوصول بكفاءة إلى أنماط يصعب ملاحظتها، لكنه يعتمد على وعي الشخص بذاته، وتفسيره، ودافعيته، وأسلوب استجابته. أما أساليب الأداء المبيّن فقد تكون أقرب إلى سلوك العمل، لكنها غالباً أعلى تكلفة وأضيق نطاقاً وتتأثر بجودة المقيم أو المحاكاة. النظام القابل للدفاع يطابق الأسلوب مع السمة، ولا يفترض أن أسلوباً واحداً متفوق دائماً.

عائلة الأسلوب	أمثلة	الأنسب له	سوء الاستخدام الشائع
القدرة المعرفية والاستعداد	الاستدلال اللفظي والعدي والمجرد والمكاني والميكانيكي.	التعلم، وحل المشكلات، والتعامل مع التعقيد عندما يكون ذا صلة بالوظيفة.	افتراض أن درجة عامة واحدة تفسر الأداء كله؛ تجاهل إمكانية الوصول ونتائج المجموعات الفرعية.
اختبارات المعرفة والمهارة	المعرفة الفنية، واللغة، والبرمجة، والجداول الإلكترونية، والسلامة.	مستوى الكفاءة الحالي في مجال محدد.	اختبار معلومات هامشية أو محتوى قديم أو معرفة يسهل تعلمها بعد التعيين.
عينات العمل والمحاكاة	سلة الأعمال، ولعب الأدوار، والعروض، ومهام البرمجة، وتحليل الحالات.	السلوك الوظيفي القابل للملاحظة والمهارات التطبيقية.	تعليمات أو تسجيل غير موحد تكلفيات غير مدفوعة تتجاوز المعقول.
اختبارات الحكم الموقفي	مواقف عمل مكتوبة أو بالفيديو أو المحادثة.	الحكم في مواقف مرتبطة بالعمل.	مفاتيح تصحيح غامضة، أو تحميل ثقافي، أو بدائل يكون «الصحيح» فيها واضحاً بشكل مفرط.
المقابلات المنظمة	أسئلة سلوكية وموقفية وفنية مع تسجيل قائم على مؤشرات محددة.	الخبرة، والحكم، والدافعية، والأدلة التفاعلية.	اعتبار المقابلة منظمة لمجرد كتابة الأسئلة مسبقاً.
الشخصية وأسلوب العمل	قوائم السمات ومقاييس أسلوب العمل الضيقة.	السلوك المعتاد، والتطوير، وفرضيات الفريق والدور.	التصنيفات الحتمية، أو الاستنتاج السريبي، أو حدود قرار صارمة غير مدعومة.
القيم والدوافع والاهتمامات	الاهتمامات المهنية، والتفضيلات الدافعية، والقيم.	التطوير، والإرشاد المهني، وحوارات الارتباط الوظيفي.	مساواة التشابه مع الثقافة الحالية بالجدارة.
النزاهة والبيانات السيرية	النزاهة الصريحة، والنزاهة المرتبطة بالشخصية، وبيانات الخبرة.	معايير محددة عندما يكون الاستخدام قانونياً ومتحققاً منه بالشكل المناسب.	أسئلة متطفلة، أو مفاتيح تجريبية غامضة، أو مؤشرات تعكس فرصاً تاريخية غير متكافئة.
التقييم متعدد المقيمين والاستبيانات التنظيمية	360 درجة، والارتباط الوظيفي، والثقافة، ومناخ الفريق.	التطوير والتشخيص التنظيمي.	استخدام تقييمات تطويرية سرية في الاختيار أو الإجراءات التأديبية.

الغرض يغيّر متطلبات الدليل

الغرض	السؤال الأساسي	تركيز الأدلة
الاختيار	من الأكثر احتمالاً لتلبية متطلبات الأداء المحددة؟	الصلة بالوظيفة، والتنبيؤ، والتوحيد، والعدالة، وأخطاء التصنيف، وردود فعل المتقدمين.
الترقية أو التعاقب	من الجاهز الآن، أو لاحقاً، أو الأنسب لمستوى مختلف؟	بُنى خاصة بالمستوى، وفرصة إظهار القدرة، والتميز بين الإمكانيات والأداء.
التطوير	ما نقاط القوة والمخاطر وأولويات التعلم التي ينبغي استكشافها؟	وضوح البنية، والثبات، والفائدة التطويرية، وجودة التغذية الراجعة، والسلامة النفسية.
الإرشاد المهني	ما البيئات والأنشطة التي قد تلائم الشخص؟	الاهتمامات، والقيم، ومفهوم الذات، والخيارات، والتفسير غير الحتمي.
العمل مع الفرق أو القيادات	كيف قد تؤثر الأنماط في التفاعل والفاعلية؟	لغة مشتركة، والسرية، ودليل يدعم الاستخدام على مستوى الفريق، وجودة التيسير.
بحوث الموظفين	ما مدى انتشار الخبرات أو الاتجاهات وما أنماطها؟	أخذ العينات، وإخفاء الهوية، وتكافؤ القياس، ومعدلات الاستجابة، والتجميع، وقابلية النتائج للتحويل إلى فعل.

لا تعد استخدام الأداة في أغراض مختلفة بالافتراض

الأداة التي تدعم حواراً تدريبياً ليست بالضرورة مناسبة لاستبعاد المتقدمين. الاختيار يطرح أسئلة مختلفة: هل يستطيع المستجيبون إدارة انطباعاتهم؟ هل نفهم الفروق بين المجموعات؟ هل المقاييس مرتبطة بالدور؟ هل قاعدة القرار متحقق منها؟ وهل يمكن الدفاع عن النتائج؟ وبالمثل قد يكون استبيان الارتباط الوظيفي ثابتاً عند الإبلاغ على مستوى المجموعة، لكنه غير مناسب لقرارات الأداء الفردية. يجب أن يتبع الدليل الاستخدام المقصود.

الأسلوب يتبع البنية؛ والدليل يتبع الاستخدام؛ والحوكمة تتبع مستوى المخاطر.

مبدأ insyt2:
الأسلوب يتبع
البنية

مراكز التقييم عمليات وليست مباني

مركز التقييم هو عملية موحدة متعددة الأساليب، يلاحظ فيها مقيمون مدربون المشاركون عبر تمارين مرتبطة بالوظيفة، ثم يدمجون الأدلة وفق أبعاد محددة. يوم من المقابلات في موقع فعلي لا يُعد بالضرورة مركز تقييم. تعتمد الجودة على تصميم التمارين، والتغطية، وتدريب المقيمين، والملاحظة السلوكية، ومؤشرات التسجيل، وقواعد الدمج، والمراقبة.

الفصل

04

مستوى المخاطر والتناسب

كلما زادت أهمية القرار وعواقبه، وجب أن تكون الأدلة والضوابط أقوى.

لماذا يهم تصنيف المخاطر؟

لن تقوم insyt2 أبداً بـ...

تطبيق معيار الحوكمة نفسه على استبيان نبض منخفض المخاطر وعلى قرار تعاقب قيادي عالي المخاطر.

في insyt2 نؤمن بأن...

يمكن للاستبيان نفسه المكون من عشر مفردات أن يكون منخفض المخاطر عندما يُستخدم للتأمل الشخصي، وعالي المخاطر عندما يُستخدم لاستبعاد المتقدمين. تتحدد المخاطر وفق القرار، والفئة المتأثرة، والحجم، وقابلية العكس، والأثمة، وحساسية البيانات، وإمكانية الوصول، وحدثة المنهج، وإمكانية الاعتراض على النتيجة. يمنع التصنيف الرسمي للمخاطر انتقال ممارسات مريحة منخفضة المخاطر إلى قرارات عالية المخاطر من دون مراجعة.

الفئة	أمثلة على الاستخدام	الحد الأدنى للحوكمة
الفئة 1 - منخفضة	تأمل ذاتي أو تعلم تطوعي دون قرار فردي.	غرض وحدود واضحة، وأساسيات الخصوصية وإمكانية الوصول، وعدم إعادة الاستخدام بصورة خفية.
الفئة 2 - متوسطة	تخطيط تطويري، أو ورشة فريق، أو حوار مواهب داخلي مع مراجعة بشرية.	تفسير مؤهل، وسرية، وأدلة مناسبة، وضوابط للميسر، وحدود موثقة.
الفئة 3 - عالية	فرز المتقدمين، أو الترقية، أو تصنيف التعاقب، أو الاعتماد.	تحليل وظيفة، واستراتيجية تحقق، وتوحيد، وتيسيرات، وتحليل مجموعات فرعية، وقواعد قرار، ومراقبة، وآلية اعتراض.
الفئة 4 - حرجة أو مبتكرة	قرارات مؤتمتة بدرجة كبيرة، أو استنتاجات حساسة، أو فرز واسع النطاق بالذكاء الاصطناعي، أو استشعار سلوكي.	مساءلة تنفيذية، ومراجعة متعددة التخصصات، وتقييم أثر، وتحجّ مستقل، وإشراف بشري قوي، وتدقيق، واستجابة للحوادث.

فحص عملي للمخاطر

قيّم كل سؤال من 0 عندما لا توجد مخاطر جوهرية إلى 2 عندما توجد مخاطر جوهرية. المجموع أداة توجيه للحوكمة وليس مقياساً سيكمترياً متحققاً منه. صعد المراجعة إذا كان أي خطر منفرد شديداً حتى وإن كان المجموع الكلي متوسطاً.

سؤال المخاطر	0	1	2
هل يمكن للنتيجة أن تمنع التوظيف أو تؤخره أو تغير الدخل بصورة جوهرية؟	لا	بشكل غير مباشر	بشكل مباشر
هل يصعب عكس القرار أو الاعتراض عليه؟	سهل	متوسط	صعب
هل الفئة كبيرة أو يحتمل أن تكون عرضة للمخاطر؟	صغيرة	متوسطة	كبيرة أو عرضة للمخاطر
هل تستخدم العملية بيانات حساسة، أو مراقبة، أو قياسات حيوية، أو سمات مستنتجة؟	لا	محدود	جوهري
هل التسجيل مؤتمت، أو غامض، أو تكيفي، أو كثير التحديث؟	لا	جزئياً	بدرجة كبيرة
هل يمكن للإعاقة أو اللغة أو الثقافة أو الوصول التقني أن تؤثر في الأداء؟	منخفض	محتمل	مرجح
هل أدلة الصدق المحلية أو أدلة المجموعات الفرعية محدودة؟	قوية	مختلطة	ضعيفة
هل قد يؤدي الخطأ إلى ضرر قانوني أو متعلق بالسلامة أو السمعة أو الأفراد؟	منخفض	متوسط	عالٍ

تصعيد insyt2:
يجب أن تغيّر
المخاطر مستوى
الحكمة

المخاطر العالية، والبيانات الجديدة، والتسجيل الغامض، والحجم الكبير، ومحدودية إمكانية الاعتراض، أو الفئات المعرضة للمخاطر يجب أن تطلق مراجعة متعددة التخصصات - لا مجرد نموذج موافقة أطول.

أسئلة التناسب

- هل السمة مهمة بما يكفي لتبرير قياسها؟
- هل الأسلوب أقل تدخلاً من البدائل المتاحة؟
- هل عبء المشارك متناسب مع قيمة القرار؟
- هل دقة الدرجة تبرر درجة التفصيل في القرار؟
- هل جمع البيانات والاحتفاظ بها محدودان بما هو ضروري؟
- هل تستطيع المراجعة البشرية الهادفة اكتشاف الأخطاء وتصحيحها؟
- هل المنفعة التجارية معقولة ومقاسة بدلاً من افتراضها؟

ورقة عمل المخاطر والتناسب

القرار والفئة المتأثرة

الاستخدام المقصود وفئة المخاطر

فجوات الأدلة والافتراضات

مخاطر إمكانية الوصول والمجموعات الفرعية

مسار المراجعة البشرية والاعتراض

مالك المراقبة وتاريخ المراجعة

ملاحظة عمل: سجّل مصادر الأدلة، والافتراضات، والقيود، والمسؤولين، وتاريخ المراجعة.

علم القياس

المفاهيم التي تفصل التقييم القائم على الأدلة عن مجرد «مسرحة الدرجات».



في هذا الجزء

05 البنى، وتحليل الوظيفة، ومخططات التقييم

06 الثبات، والدقة، وعدم اليقين في الدرجات

07 الصدق والتحقق

08 المعايير، والمقاييس، والمفردات، والعدالة، وحدود القرار

العلم • العدالة • جودة القرار • الحكم المهني

رؤية insyt2

السؤال ليس أي اختبار يبدو أكثر إبهاراً، بل أي دليل يخدم القرار، والأشخاص المتأثرين به، وما يحدث بعد ذلك بصورة أفضل.

الفصل

05

البُنى، وتحليل الوظيفة، ومخططات التقييم

قبل قياس الأشخاص، عرّف السمة وعلاقتها بالعمل.

البنية فكرة تفسيرية

في insyt2 نؤمن بأن...

لن تقوم insyt2 أبداً بـ...

السماح لاسم يقدمه موزّد أو لُبعد ذي علامة تجارية بأن يحل محل تعريف واضح للبنية.

من العمل إلى الدليل: سلسلة التقييم

الخطوة	السؤال	المخرج
1. النتيجة التجارية	ما الذي يجب على الدور أو البرنامج تحقيقه؟	نتائج الأداء، أو السلامة، أو الخدمة، أو التعلم، أو الاحتفاظ، أو التطوير.
2. تحليل العمل	ما المهام والسياقات والقيود والتفاعلات التي تنتج تلك النتائج؟	خريطة المهام والسياق، والحوادث الحرجة، والفروق بين المستويات.
3. فرضية السمة	ما المعارف والمهارات والقدرات والخصائص الأخرى التي تمكن الأداء؟	بُنى محددة مع مؤشرات سلوكية.
4. نموذج الدليل	ما الاستجابة القابلة للملاحظة التي ستدعم كل استنتاج؟	سلوكيات، ومنتجات عمل، وإجابات، وتقديرات، وعمليات استجابة.
5. مخطط التقييم	ما مقدار المحتوى وأي أسلوب سيأخذ عينة من كل مجال؟	مصفوفة تغطية، ومواصفات مفردات، وتمارين، وخطة تسجيل.
6. حجة التحقق	ما الأدلة التي تبرر التفسير والاستخدام المقصودين؟	خطة بحث ومراقبة.

تحليل الوظيفة لأغراض التقييم

لكل بنية مقترحة، اجمع أدلة على الأهمية، والتكرار أو فرص الممارسة، والمستوى المطلوب، وقابلية التدريب، ومدى تميزها عن غيرها. قد تكون سمة ما مهمة لكنها غير مناسبة للاختبار قبل التوظيف لأنها تُتعلم بسرعة، أو لا تتاح للمتقدمين فرص سابقة متكافئة لإظهارها، أو يمكن ملاحظتها بصورة أفضل من خلال عينة عمل. وقد يستحق مطلب قليل التكرار وزناً كبيراً إذا كانت عواقب الفشل فيه شديدة.

نقص البنية والتلوث

المشكلة	المعنى	المثال	الضبط
نقص البنية	يأخذ التقييم عينة ضيقة أكثر من اللازم من المجال المقصود.	اختبار قيادة يقيس الثقة في العرض لكنه يهمل الحكم، والشمول، والتنفيذ، والأخلاقيات.	ضع مخططاً للمجال، واستخدم أساليب متعددة، وراجع التغطية مع خبراء متنوعين.
تباين غير ذي صلة بالبنية	تتأثر الدرجات بخصائص خارج البنية المقصودة.	اختبار عددي يستخدم لغة معقدة بلا ضرورة، فيضيف عبئاً على القراءة.	أزل المتطلبات غير الضرورية، وراجع إمكانية الوصول، وادرس أنماط المجموعات الفرعية.
نقص المعيار	مقياس النتيجة يفوت جوانب مهمة من الأداء.	تقييمات المدير لا تشمل العمل الجماعي وسلوك السلامة.	استخدم نموذجاً للمعيار ومؤشرات متعددة مضبوطة الجودة.
تلوث المعيار	تتضمن النتيجة تأثيرات غير ذات صلة.	نتائج المبيعات تعكس جودة المنطقة أكثر من قدرة الفرد.	اضبط السياق واستخدم تقديرات سلوكية وفترات متعددة.

مثال لمخطط: مدير خدمة عملاء في المستوى الإشرافي الأول

البنية أو المجال	دليل العمل	الأسلوب	التغطية
تحديد الأولويات التشغيلية	يدير قوائم الانتظار والموارد والتصعيد ومستويات الخدمة.	تمرين سلة أعمال محدد بزمان مع متابعة في مقابلة منظمة.	30%
الحكم التدريبي	يشخص الأداء ويجري حوارات عادلة ومحددة.	اختبار حكم موقف في الفيديو مع لعب دور.	25%
تفسير البيانات	يستخدم مؤشرات الخدمة ويحدد الأسباب الجذرية.	تمرين عددي في سياق العمل.	20%
الحكم تجاه العميل والمخاطر	يوازن بين الخدمة والسياسة والهشاشة والتصعيد.	اختبار حكم موقف في تحليل حالة.	15%
التعلم والتكيف	يحدث الممارسة بعد تغيير السياسات والتكنولوجيا.	مقابلة سلوكية منظمة.	10%

تحليل الوظيفة من المهارات ذات القيمة العالية لمن يبدأ في المجال. تحويل الحوادث الحرجة وأدلة الدور إلى مخطط تقييم هو بالفعل عمل تخصصي.

منظور insyt2 للمسيرة المهنية: ابدأ بتحليل الوظيفة

تجنب «تسويق» البنى

قد تجمع التسميات التجارية مثل الرشاقة، أو الإمكانات، أو العقلية الرقمية، أو الحضور التنفيذي عدة بُنى معروفة أو تخفي تعريفاً غير واضح. اطلب من المورد ربط التسمية بسلوك قابل للملاحظة، ونظرية، ومحتوى المفردات أو التمارين، والبنية العاملة، ودليل المعيار، والحدود. الاسم الجديد لا يخلق سمة نفسية جديدة.

الفصل

06

الثبات، والدقة، وعدم اليقين في الدرجات

القياس المتسق أكبر من مجرد معامل واحد.

ما معنى الثبات؟

في insyt2 نؤمن بأن...

لن نقوم insyt2 أبداً بـ...

التعامل مع فروق صغيرة في الدرجات كفروق ذات معنى في ترتيب الأشخاص من دون النظر إلى خطأ القياس.

الثبات هو اتساق الدرجات أو دقتها تحت شروط محددة. وهو يسأل: إلى أي مدى يعكس التباين في الدرجات المرصودة فروقاً مستقرة نسبياً في السمة المقصودة، بدلاً من التأثيرات العشوائية أو غير المرغوبة؟ الثبات ليس خاصية «موجودة أو غير موجودة»، وليس صفة عالمية ثابتة للأداة. فهو يعتمد على الفئة، والغرض، والدرجة، وطريقة الإدارة، وطول الأداة، وجودة المفردات، وسلوك المقيمين، والنموذج المستخدم.

المعامل المرتفع لا يثبت أحادية البعد، أو الصلة بالعمل، أو العدالة، أو الصدق. فقد تنتج مجموعة من المفردات شبه المتطابقة اتساقاً داخلياً مرتفعاً مع أخذ عينة شديدة الضيق من المجال. وبالمقابل قد يكون لمركب واسع صُمم عمداً اتساق داخلي متوسط لأنه يمثل عدة جوانب مهمة. فسر الثبات في ضوء تصميم الدرجة ومستوى المخاطر المرتبط بالقرار.

نوع الدليل	السؤال الذي يجب عنه	الاستخدامات الشائعة	ما يجب الانتباه إليه
الاتساق الداخلي	هل تتسق المفردات التي يُفترض أن تسهم في الدرجة؟	المقاييس متعددة المفردات، ألفا، أوميغا، والتجزئة النصفية.	ليس اختصاراً للاستقرار أو أحادية البعد؛ يتأثر بالطول والتكرار.
إعادة الاختبار	هل تبقى الدرجات مستقرة عبر فاصل زمني مناسب؟	السمات والمعرفة والقدرة عندما لا يُتوقع تغير حقيقي.	الذاكرة، والتعلم، والتغير الحقيقي، وطول الفاصل، والتسرب من العينة.
النماذج البديلة	هل تنتج النماذج المكافئة درجات قابلة للمقارنة؟	الاختبارات الآمنة، وإعادة الاختبار، وبنوك المفردات.	قد تختلف النماذج في المحتوى أو الصعوبة وقد يلزم الربط.
الاتفاق بين المقيمين	هل يصدر المقيمون المدربون أحكاماً متقاربة؟	المقابلات، والمحاكاة، والعمل المكتوب، ومراكز التقييم.	قد يتعايش الاتفاق مع تحيز مشترك؛ مؤشرات التسجيل والمعايرة ضرورية.
اتساق القرار	هل يحصل الأشخاص على التصنيف نفسه عند التكرار؟	النجاح/الرسوب، والاعتماد، وحدود القرار.	غالباً يكون الأضعف قرب العتبة حتى عندما يبدو الثبات الكلي مقبولاً.
الدقة الشرطية	ما مدى دقة القياس عند مستويات مختلفة من الدرجة؟	نظرية الاستجابة للمفردة، والاختبار التكيفي، والمناطق حول الحدود.	قد يخفي متوسط الثبات ضعفاً في الدقة حيث تُتخذ القرارات.

الاتساق الداخلي: ألفا وما بعدها

يُبلغ عن معامل ألفا كرونباخ على نطاق واسع، لكنه يعتمد على افتراضات كثيراً ما تُهمل. يزداد مع طول الاختبار وقد يرتفع بسبب التكرار غير المفيد في المحتوى. قد يكون معامل أوميغا أنسب عندما تختلف تشبعات المفردات، لكنه يعتمد أيضاً على نموذج عاملي مناسب. لا يوجد معامل عالمي يجعل مقياساً مقبولاً في كل استخدام. عادة ما تتطلب القرارات الفردية عالية المخاطر دقة أقوى من البحوث الاستكشافية على مستوى المجموعات، وينبغي فحص الدقة حول عتبات القرار.

لن تختزل insyt2 الثبات إلى رقم واحد

لا تطبق قاعدة عامة من نوع 70 من دون النظر إلى الغرض، واتساع البنية، وملاءمة النموذج، ومستوى الدرجة، والعينة، والعواقب.

الخطأ المعياري للقياس

يصبح الثبات أكثر فائدة عندما نترجمه إلى عدم يقين في الدرجة. في النموذج الكلاسيكي يقدّر الخطأ المعياري للقياس الانتشار المعتاد للدرجات المرصودة حول «الدرجة الحقيقية» المفاهيمية للشخص، بوحدات المقياس نفسه.

$$SEM = SD \times \sqrt{1 - \text{reliability}}$$

حيث SD هو الانحراف المعياري للدرجات في الفئة ذات الصلة، ويجب أن يتوافق تقدير الثبات مع الدرجة والاستخدام.

مثال: مقياس T له انحراف معياري $SD = 10$ وثبات $= .84$. إذن $SEM = 10 \times \sqrt{.16} = 4$. ويمكن تقريب فاصل 95% البسيط إلى الدرجة المرصودة $\pm SEM \times 1.96$. إذا كانت الدرجة المرصودة 62، يكون النطاق تقريباً من 54 إلى 70. هذا النطاق واسع بما يكفي لإظهار لماذا لا ينبغي اعتبار فارق نقطتين فرقاً ذا معنى في ترتيب الأشخاص. قد تنتج النماذج الأكثر تقدماً فواصل غير متماثلة أو خاصة بكل مستوى من الدرجات؛ استخدم الأسلوب الذي يوصي به الدليل الفني.

فاصل 95% تقريبي = الدرجة المرصودة $\pm SEM \times 1.96$

هذه صيغة تفسيرية تقريبية، وليست بديلاً عن نموذج التسجيل والدليل الفني الخاص بالأداة.

كيف تؤثر الدقة في القرارات؟

- استخدم فترات الثقة عند تفسير الدرجات الفردية، وخصوصاً قرب العتبات.
- تجنب المبالغة في ترتيب الأشخاص عندما تتداخل فترات الثقة بصورة كبيرة.
- في المحاكاة، أبلغ عن دقة المقيمين والتمارين، لا عن المركب الكلي فقط.
- ادرس اتساق القرارات حول حدود القرار، وضع قواعد لإعادة الاختبار أو المراجعة عند الحاجة.
- لا تعرض منازل عشرية أكثر مما يسمح به الدليل.
- أعد فحص الثبات بعد الترجمة، أو تغيير المنصة، أو تغيير الوقت، أو استخدام الأداة مع فئة جديدة.

**مقياس insyt2
للتقارير: طابق
اللغة مع مستوى
الدقة**

يجب أن تتناسب لغة التقرير مع قوة الدليل. عبارات مثل «تشير النتائج إلى»، و«تتسق مع»، و«تقع ضمن النطاق فوق المتوسط» غالباً أكثر قابلية للدفاع من التصنيفات القطعية.

07

الفصل

الصدق والتحقق

السؤال المركزي هو ما إذا كانت الأدلة تدعم المعنى والاستخدام المقترحين للدرجات.

الصدق حجة تدعمها الأدلة

في insyt2 نؤمن بأن...

لن نقوم insyt2 أبداً بـ...

استخدام عبارة «تم التحقق منه» من دون بيان الفئة، والغرض، والأدلة، والقيود التي تدعم الادعاء.

تركز المعايير الحديثة للاختبارات على صدق تفسيرات الدرجات واستخداماتها المقصودة. والسؤال ليس ببساطة: «هل هذا الاختبار صادق؟» بل: «هل التفسير المقترح واستخدامه في القرار مدعومان بما يكفي لهذه الفئة وهذا الغرض وهذا السياق؟» تعتمد الإجابة على حجة مترابطة تصل تعريف البنية بالمحتوى، وعمليات الاستجابة، والبنية الداخلية، والعلاقات بمتغيرات أخرى، والعدالة، والعواقب.

الصدق تراكمي ومشروط. فالدليل الذي يدعم استخدام درجة استدلال في توظيف الخريجين لا يدعم تلقائياً استخدام الدرجة نفسها للترقية إلى عائلة وظيفية مختلفة، أو في نسخة مترجمة، أو في إدارة غير مراقبة عبر الهاتف المحمول، أو لاستنتاج سريري. قد تؤدي التغييرات في المحتوى أو التسجيل أو أسلوب التقديم أو الفئة أو قواعد القرار إلى تغيير الاستنتاج وتستدعي تحقّقاً إضافياً.

خمس مصادر شائعة لأدلة الصدق

مصدر الدليل	الأسئلة	أمثلة
محتوى الاختبار	هل يمثل المحتوى المجال المحدد بالمستوى المناسب ويتجنب المواد غير ذات الصلة؟	تغطية المخطط، وتقديرات الخبراء، وربط المحتوى، ومراجعة المفردات.
عمليات الاستجابة	هل يخطر الأشخاص في العمليات المعرفية أو السلوكية التي تفترضها الدرجة؟	دراسات التفكير بصوت عالٍ، وأنماط زمن الاستجابة، وإدراك المقيمين، وبحوث قابلية الاستخدام.
البنية الداخلية	هل تتوافق العلاقات بين المفردات أو التمارين أو المقيمين مع نموذج الدرجة المقترح؟	التحليل العاملي، والأبعاد، والاعتماد المحلي، والثبات، والتكافؤ.
العلاقات مع متغيرات أخرى	هل ترتبط الدرجات بالمعايير والبُنى الأخرى كما تتنبأ النظرية؟	الصدق التنبؤي والمتزامن، والأدلة التقاربية والتمييزية، ومقارنات المجموعات.
العواقب والاستخدام	ما الآثار المقصودة وغير المقصودة الناتجة عن الاستخدام وقواعد القرار؟	المنفعة، ونتائج المجموعات الفرعية، وردود فعل المتقدمين، وأخطاء التصنيف، والتحويل، والسلوك التنظيمي.

الأدلة المرتبطة بالمحتوى

تزداد أهمية أدلة المحتوى في اختبارات المعرفة، وعينات العمل، والمحاكاة، والمقابلات المنظمة، والمقاييس شديدة الارتباط بوظيفة معينة. العملية القابلة للدفاع تعرّف مجال العمل، والبنى، والمستوى، وخطة أخذ العينة قبل كتابة المفردات. يجب أن يقيم خبراء الموضوع الصلة والتمثيل باستخدام معايير واضحة، وأن توثق الاختلافات بدلاً من إخفائها بمتوسط حسابي من دون نقاش. دليل المحتوى لا يعني أن المفردة تذكر الوظيفة فقط؛ بل يجب أن تستثير دليلاً ذا صلة بالمتطلب المستهدف.

الأدلة المرتبطة بالمعيار

تدرس دراسات التحقق المرتبط بالمعيار العلاقات بين درجات التقييم والنتائج مثل الأداء الوظيفي، والنجاح في التدريب، والسلامة، وجودة المبيعات، والتقدم الوظيفي، أو الاحتفاظ بالموظفين. تجمع التصميم التنبؤية درجات التقييم قبل فترة قياس المعيار، وعادة ما تحجب هذه الدرجات عن صانعي القرار؛ وهي الأقرب إلى واقع الاختيار لكنها تستغرق وقتاً. أما التصميم المتزامنة فتقيم الموظفين الحاليين وتربط درجاتهم بالمعايير الحالية؛ وهي أسرع لكنها معرضة لتقييد المدى، وتأثيرات البقاء، والخبرة، وتلوث المعيار.

الارتباط لا يفسر نفسه. قيم حجم العينة وعدم اليقين، وجودة المعيار، والتصميم، وتقييد المدى، والبيانات المفقودة، وأنماط المجموعات الفرعية، والتحقق المتقاطع، والأهمية العملية. وانتبه إلى تسرب البيانات: فقد تبدو الخوارزمية تنبؤية عندما تتسلل معلومات عن النتيجة أو متغيرات ما بعد التعيين إلى الخصائص المدخلة. قيم الصدق الإضافي - أي ما إذا كان الأسلوب يضيف أدلة مفيدة فوق الأساليب القائمة - بدلاً من الاحتفاء بعلاقة منفردة تكرر معلومات متاحة بالفعل.

الأدلة المرتبطة بالبنية

تفحص أدلة البنية ما إذا كانت الدرجات تتصرف كما هو متوقع داخل شبكة من النظرية والنتائج. ينبغي لمقياس الضمير المهني مثلاً أن يرتبط إيجابياً بالمقاييس المعروفة للضمير المهني وبالسلوك الوظيفي ذي الصلة، ولكن لا ينبغي أن يرتبط بقوة شديدة ببنى غير مرتبطة بحيث يبدو وكأنه يقيس كل شيء. يمكن أن تسهم البنية العالمية، والعلاقات التقاربية والتمييزية، وفرضيات المجموعات المعروفة، والتغير التطوري، وعمليات الاستجابة. والتحقق من البنية عملية مستمرة؛ ولا يكتمل بمجرد تسمية عامل بعد فحص المفردات.

حجة تحقق في الممارسة

الادعاء	الأدلة المطلوبة	التحديات التي يجب فحصها
الدرجة تمثل الحكم في مخاطر العميل.	تعريف البنية، ومخطط المواقف، ومراجعة الخبراء، ودراسة عمليات الاستجابة، والبنية الداخلية.	عبء القراءة، ومعرفة السياسات، والرغبة الاجتماعية، ومفاتيح التصحيح الغامضة.
الدرجات الأعلى تشير إلى حكم مستقبلي أفضل.	علاقة تنبؤية بمعيار مضبوط الجودة، والتحقق المتقاطع، والقيمة الإضافية.	تحيز المعيار، وصغر العينة، والتسرب، وتقييد المدى، وعدم استقرار النموذج.
يمكن للدرجة دعم عتبة فرز.	الدقة حول العتبة، ومنطق حد القرار، وتكاليف الخطأ، ونتائج المجموعات الفرعية.	قرارات حادة مبنية على درجات غير دقيقة وأثر سلبي غير ضروري.
الاستخدام عادل ومتاح.	اختبار إمكانية الوصول، وبروتوكول التيسيرات، وتحليل DIF أو التكافؤ، والتنبؤ حسب المجموعات، والنتائج، وبحوث المرشحين.	عوائق غير مرتبطة بالبنية، ومؤشرات بديلة حساسة، ومنصة غير متاحة، وفروق في فرص الاستعداد.
يبقى النظام صالحاً بعد الإطلاق.	ضبط الإصدارات، ومراقبة الانحراف، وسجلات الحوادث، وإعادة التحقق دورياً.	انكشاف المفردات، والتدريب على الاختبار، وتغير الفئة، وتغير المعيار، وتحديثات النموذج.

لا تجدد ترتيب الأساليب

تقدم بحوث الاختيار أدلة عامة مهمة، لكن لا يوجد أسلوب له معامل صدق ثابت لا يتغير. تعتمد تقديرات التحليل التلوي على جودة الدراسات، والمعايير، والوظائف، والفئات، والتصحيحات، والفترة الزمنية. وقد أظهرت مراجعات حديثة لبعض التقديرات الكلاسيكية أن بعض القيم المتداولة على نطاق واسع تعرضت لتصحيحات مبالغ فيها، وأن الاستنتاجات يجب أن تتحدث مع تحسن الأدلة. استخدم البحوث لتكوين الفرضيات وتصميم الأنظمة، ثم قيم الملاءمة والدليل المحلي بدلاً من نسخ جدول ترتيب عالمي.

انضباط insyt2 في الأدلة: شريحة مبيعات ليست برنامج تحقق

نطلب الفرضية، والتصميم، والعينة، والمعيار، وعدم اليقين، ونتائج المجموعات الفرعية، والتحقق المتقاطع، وعواقب القرار - لا مجرد معامل ارتباط رئيسي.

الفصل

08

المعايير، والمقاييس، والمفردات، والعدالة، وحدود القرار

تصبح الدرجة الخام مفيدة فقط عندما توضع داخل إطار مرجعي وقاعدة قرار يمكن الدفاع عنها.

الدرجات الخام والمقاييس الشائعة

في insyt2 نؤمن بأن...

لن نقوم insyt2 أبداً بـ...

عرض المئينات أو الأشرطة اللونية كحقائق تفسر نفسها من دون وصف المجموعة المعيارية، وعدم اليقين، والعواقب.

قد تكون الدرجة الخام عدد الإجابات الصحيحة، أو مجموع التقديرات، أو متوسط درجات المقيمين، أو ناتج نموذج. وهي بمفردها نادراً ما تخبر المستخدم ما إذا كان الأداء قوياً أو ضعيفاً أو معتاداً أو كافياً. قد يُفسر الأداء بمقارنته بمجموعة معيارية، أو معيار محدد، أو درجة سابقة، أو مقياس نظري. يجب أن يتناسب إطار المرجعية المختار مع القرار.

المقياس	التعريف	تنبيه عند التفسير
النسبة المئوية الصحيحة	عدد الإجابات الصحيحة ÷ إجمالي المفردات القابلة للتسجيل.	تعتمد الصعوبة على الاختبار؛ 80% في اختبار لا تساوي بالضرورة 80% في اختبار آخر.
الرتبة المئينية	نسبة أفراد المجموعة المرجعية الذين حصلوا على الدرجة نفسها أو أقل.	المئينات ترتيبية والمسافات بينها غير متساوية.
درجة Z	الدرجة الخام ناقص المتوسط، مقسوماً على الانحراف المعياري.	القيمة السالبة لا تعني «سلباً»؛ المعنى يعتمد على التوزيع المرجعي.
درجة T	$z \times 10 + 50$	المقياس المريح لا يلغي خطأ القياس أو عدم اعتدال التوزيع.
Sten	عادة نطاق معياري من 1 إلى 10.	الفئات الواسعة تفقد التفاصيل ويبالغ بسهولة في تفسير حدودها.
Stanine	عادة نطاق معياري من 1 إلى 9.	صُمم للتصنيف الواسع، لا للترتيب الدقيق.
مستوى مرجعي للمعيار	الأداء بالنسبة إلى معيار أو مجال محدد.	يتطلب وضع معيار يمكن الدفاع عنه ومهماً ممثلة للمجال.
Theta في IRT أو درجة محولة	تقدير قائم على نموذج على متصل كامن.	المعنى يعتمد على ملاءمة النموذج، والمعايرة، والربط، والدقة الشرطية.

$$T = 50 + 10z \text{ و } z = (X - M) / SD$$

حيث X الدرجة الخام، M متوسط المجموعة المرجعية، وSD انحرافها المعياري.

جودة المجموعة المعيارية

ينبغي أن تكون المجموعة المعيارية ذات صلة بالمقارنة المقصودة، وكبيرة بما يكفي، ومأخوذة بعينة مناسبة، وموصوفة بشفافية، وحديثة بصورة معقولة. عبارة «مهيئون عالميون» ليست وصفاً نياً. أسأل من شملتهم العينة، ومن أي دول ولغات، وكيف جرى استقطابهم، وهل كانوا متقدمين أم موظفين، ومستوياتهم الوظيفية، والقطاعات، والتعليم، والتواريخ، والأوزان، والاستبعادات، وكيف عولجت البيانات المفقودة. يمكن للدرجة الخام نفسها أن تنتج مئينات مختلفة في معايير الخريجين أو المديرين أو التنفيذيين أو المعايير الوطنية أو معايير المتقدمين. استخدم المعايير الواسعة عندما يتطلب القرار فعلاً مقارنة بسوق عمل واسع.

- استخدم معايير الدور أو المستوى أو القطاع فقط عندما تكون ذات صلة، وليس عندما تكون ضيقة إلى حد يعيد إنتاج التجانس الحالي.
- تجنب استخدام معايير الموظفين الحاليين المحليين بوصفها تعريفاً غير قابل للنقاش لمواهب المستقبل.
- أعد بناء المعايير أو افحصها عندما تتغير الفئات أو الاستعداد أو نمط تقديم الاختبار أو اللغة أو انكشاف المفردات بصورة جوهرية.
- وثق إصدار المعايير المستخدم في كل تقرير للدرجات.

التفسير المعياري، والتفسير المرجعي للمعيار، والتفسير الإجمالي

يسأل التفسير المعياري: كيف يقارن الشخص بالآخرين؟ بينما يسأل التفسير المرجعي للمعيار: هل أظهر الشخص المعرفة أو الأداء المحدد؟ قد يكون المرشح أعلى من المتوسط لكنه دون معيار السلامة، أو أقل من معيار شديد الانتقائية لكنه قادر تماماً على الدور. يعتمد الاختيار التنافسي غالباً على الدليل النسبي، بينما يتطلب الاعتماد أو الحد الأدنى للكفاءة معياراً معرفاً مسبقاً.

تفرض الصيغ الإيساطية مفاضلات بين البدائل، فتجعل الدرجات داخل الشخص مرتبطة جزئياً بعضها ببعض. يمكن أن تكون الملفات الإيساطية التقليدية مفيدة للتأمل، لكنها إشكالية عند مقارنة الأفراد. تستطيع بعض نماذج الاختيار القسري الحديثة أحياناً استرداد تقديرات معيارية للسمات، لكن يجب فحص النموذج، وتصميم المفردات، والتسجيل، والثبات، والصدق. الاختيار القسري ليس محصناً تلقائياً ضد إدارة الانطباع، ولا مناسباً تلقائياً للاختيار.

تحليل المفردات ونماذج القياس

الدليل	ما الذي يشير إليه	التفسير
صعوبة المفردة أو نسبة تأييدها	نسبة الإجابة الصحيحة أو توزيع الاستجابات.	استهدف النطاق المطلوب؛ قد تضيف المفردات شديدة السهولة أو الصعوبة تمييزاً قليلاً لكنها قد تغطي محتوى أساسياً.
تمييز المفردة	مدى قدرة المفردة على التفريق بين ذوي الدرجات الأعلى والأدنى.	القيم المنخفضة أو السالبة قد تشير إلى غموض أو خطأ في المفتاح أو تعدد الأبعاد أو ضعف المحتوى.
ارتباط المفردة بالمجموع	ارتباط المفردة ببقية المقياس.	مؤشر تشخيصي مفيد وليس قاعدة الاحتفاظ الوحيدة؛ احم تغطية المجال.
فاعلية البدائل المشتتة	هل تجذب البدائل الخاطئة أخطاءً معقولة؟	البدائل غير المعقولة تقلل المعلومات وتكشف الإجابة الصحيحة.
التحليل العاملي	كيف تتجمع المفردات أو المؤشرات بالنسبة إلى بنية مقترحة؟	مؤشرات الملاءمة أدلة، لا شهادة تلقائية؛ النظرية والتكرار مهمان.
نظرية الاستجابة للمفردة	احتمال الاستجابة بوصفه دالة في مستوى الشخص وخصائص المفردة.	تدعم بنوك المفردات والدقة الشرطية، لكنها تحتاج إلى ملاءمة ومعايرة وحوكمة.
الأداء التفاضلي للمفردة DIF	هل يستجيب أشخاص متقاربون في السمة بصورة مختلفة باختلاف المجموعة؟	إشارة تستدعي تحقيقاً موضوعياً، وليست دليلاً تلقائياً على التحيز.

يستكشف التحليل العاملي الاستكشافي البنية عندما تكون غير مؤكدة، بينما يختبر التحليل العاملي التوكيدي نموذجاً محدداً. يفحص تكافؤ القياس ما إذا كان المقياس يعمل بصورة متقاربة بما يكفي عبر اللغات أو المجموعات لإجراء المقارنات المقصودة. يمكن لنظرية الاستجابة للمفردة دعم بنوك المفردات، والربط، والاختبار التكيفي، ودوال المعلومات، لكنها ليست متفوقة تلقائياً على الأساليب الكلاسيكية. استخدم أبسط نموذج يدعم الاستنتاج ودورة الحياة بصورة كافية.

العدالة، والتحيز، وإمكانية الوصول

المفهوم	المعنى	ما الذي لا يثبتته
فروق متوسطات المجموعات	تختلف متوسطات الدرجات بين المجموعات في عينة ما.	لا تثبت بمفردها تحيز المفردات أو التمييز أو عدم الصلة بالوظيفة.
الأثر السلبي	تنتج العملية نتائج اختيار مختلفة جوهرياً بين المجموعات وفق القانون أو السياسة ذات الصلة.	لا يحدد بمفرده السبب ولا يكمل التحليل القانوني.
تحيز القياس	لا تمثل الدرجات البنية نفسها بصورة قابلة للمقارنة عبر المجموعات.	لا يمكن تشخيصه من فرق في المتوسط الكلي وحده.
التحيز التنبؤي	علاقة الدرجة بالنتيجة تبالغ أو تقلل التنبؤ بصورة منهجية لمجموعات معينة.	يحتاج إلى معايير ونماذج مناسبة؛ تساوي الارتباطات وحده غير كاف.
العدالة	حكم أوسع يشمل إمكانية الوصول، والمعاملة، والقياس، والتنبؤ، والقرارات، والعواقب، والقيم.	لا يمكن اختزالها في إحصائية واحدة أو في المعاملة المتطابقة.

إمكانية الوصول هي تصميم المحتوى والبيئات بحيث يستطيع أشخاص ذوو قدرات متنوعة استخدامها. أما التيسير فهو تعديل لتلبية حاجة فردية. ومن أمثلته وقت إضافي، أو فترات راحة، أو وسيلة إدخال بديلة، أو التوافق مع قارئ الشاشة، أو تكبير النص، أو قارئ أو كاتب عندما يُسمح بذلك، أو دعم لغة الإشارة، أو مكان مختلف، أو أسلوب بديل. يعتمد التعديل الصحيح على العائق وعلى البنية المقصودة. قد يكون الوقت الإضافي مناسباً عندما لا تكون السرعة جزءاً من السمة؛ أما إذا كانت السرعة جوهرية فقد نحتاج إلى طريقة بديلة لإظهار المتطلب.

مبدأ insyt2 للعدالة: صقم لوصول متكافئ

الوصول المتكافئ لا يعني دائماً شروطاً متطابقة. الهدف هو إزالة العوائق غير المرتبطة بالبنية مع الحفاظ على معنى الدرجة.

الترجمة ليست تكييفاً

يمكن للترجمة الحرفية أن تحفظ الكلمات وتغيّر في الوقت نفسه الصعوبة، والدلالة، والرغبة الاجتماعية، والملاءمة الثقافية، أو السلوك الذي تؤخذ منه العينة. يبدأ التكييف بالتأكد من أن البنية والاستخدام لهما معنى في السياق المستهدف. ويستخدم مترجمين مؤهلين، وخبراء موضوع، ومراجعة توفيقية، ومقابلات معرفية، وتجريباً، وتحليلاً للبنية، والثبات، وأداء المفردات، والمعايير، والعلاقات بالمعيار.

حدود القرار والتصنيف

حد القرار يصنف الأشخاص إلى فئات مثل ناجح/راسب، مؤهل/غير مؤهل، جاهز/غير جاهز. هذه الوضوح الظاهري يخفي عدم اليقين. قد يملك أشخاص قريبون من العتبة قدرة كامنة متشابهة لكنهم يحصلون على نتائج مختلفة بسبب خطأ القياس. وقد ترفع العتبة الأعلى متوسط الأداء المتوقع، لكنها تستبعد أشخاصاً قادرين أو تزيد الفروق بين المجموعات. العتبة المناسبة قرار سياساتي قائم على الدليل، وليست حقيقة رياضية مخبأة في التوزيع.

أساليب وضع حدود القرار

الأسلوب	كيف يعمل	الملاءمة والتنبيه
Angoff المعدل	يقدر الخبراء احتمال أن يجيب شخص ذو حد أدنى من الكفاءة عن كل مفردة بصورة صحيحة.	اختبارات المعرفة أو الاعتماد؛ يتطلب تعريفاً واضحاً للكفاءة وحكاماً مدربين.
Bookmark	يحدد الخبراء نقطة في كتيب مفردات مرتبة تمثل الحد الأدنى للكفاءة.	اختبارات مقاسة وفق IRT؛ يعتمد على معايرة المفردات وفهم الحكام.
المجموعة الحدية أو الانحدار الحدي	يستخدم أداء المرشحين الذين حُكم عليهم بأنهم على الحد وفق معيار خارجي لتحديد العتبة.	المحاكاة وتقييم الأداء المنظم؛ يجب أن تكون الأحكام الخارجية ثابتة.
المجموعات المتباينة	تقارن توزيعات مجموعات مصنفة على أنها كفوة وغير كفوة بعد.	يتطلب تصنيفات مجموعات يمكن الدفاع عنها وعينات ممثلة.
تحسين المعيار	توازن العتبة بين الحساسية، والنوعية، والمنفعة، أو تكاليف الخطأ.	سياقات التنبؤ والمخاطر؛ تجنب فرط الملاءمة وراعي العدالة والطاقة الاستيعابية.
عتبة سياساتية	تختار المؤسسة درجة بناء على الاستراتيجية أو السعة.	قد تكون مشروعة للترتيب، لكن لا ينبغي تمثيلها على أنها معيار كفاءة.

نسبة الأثر السلبي = معدل اختيار المجموعة ÷ أعلى معدل اختيار بين المجموعات
مثال: 35% ÷ 50% = 0.70. في الولايات المتحدة قد تكون هذه إشارة للمراجعة وفق قاعدة الأربعة أخماس، وليست استنتاجاً قانونياً كاملاً.

ممارسة أفضل قرب العتبة

- استخدم نطاقات درجات أو فترات ثقة بدلاً من الادعاء بأن فرق نقطة واحدة دقيق.
- أنشئ منطقة مراجعة تستخدم أسلوباً مستقلاً ذا صلة بالعمل أو مراجعة لجنة منظمة.
- حدد فترات الانتظار قبل إعادة الاختبار، والنماذج البديلة، وكيفية معالجة الحوادث التقنية.
- حلل أثر العتبات البديلة قبل التنفيذ.
- وضح ما إذا كانت العتبة تمثل الحد الأدنى للكفاءة، أو التنافس النسبي، أو السعة، أو حاجزاً للانتقال إلى مرحلة لاحقة.
- راقب اتساق القرار، ونتائج المجموعات الفرعية، وجودة التعيين، وتكاليف الأخطاء.

لن تستخدم insyt2 أبداً رقماً مريحاً لمجرد أنه سهل كحد قرار
تحتاج العتبة إلى منطق واضح لوضع المعيار، وتحليل للأخطاء، وحوكمة. الأرقام الدائرية ليست منهجية.

دورة حياة التقييم

من الاستراتيجية واختيار المورد إلى التصميم والإدارة والتسجيل والحوكمة.



في هذا الجزء

09 استراتيجية التقييم وميثاق المشروع

10 اختيار الاختبارات وتقييم الموردين

11 تطوير التقييم

12 الإدارة، وإمكانية الوصول، وتجربة المرشح

13 التسجيل، والتفسير، والتقارير، والتغذية الراجعة

14 دمج الأدلة، والتحقق، والمراقبة، وحوكمة البيانات

العلم • العدالة • جودة القرار • الحكم المهني

رؤية insyt2

السؤال ليس أي اختبار يبدو أكثر إبهاراً، بل أي دليل يخدم القرار، والأشخاص المتأثرين به، وما يحدث بعد ذلك بصورة أفضل.

09

الفصل

استراتيجية التقييم وميثاق المشروع

عرّف القرار، والدليل، والمخاطر، والملكية، ومعايير النجاح قبل التنفيذ.

ابدأ بهندسة القرار

في insyt2 نؤمن بأن...

لن نقوم insyt2 أبداً بـ...

بدء مشروع بالتسوق لشراء الاختبارات قبل أن تتفق المؤسسة على القرار الذي تريد تحسينه.

عناصر الميثاق	الأسئلة التي يجب الإجابة عنها
القرار والغرض	ما القرار الذي سيدعمه التقييم؟ هل هو استشاري، أم حاجز، أم محدد للقرار؟
الفئة والسياق	ما الأدوار والمستويات والمواقع واللغات وأنواع العقود وقنوات المرشحين الداخلة في النطاق؟
نتائج النجاح	ما النتائج التجارية، والجودة، والعدالة، والتجربة، والتشغيل التي تعرف النجاح؟
الحاجة إلى البنية والدليل	ما المعلومات ذات الصلة بالوظيفة المفقودة من العملية الحالية؟ وما الذي لن يتم تقييمه؟
فئة المخاطر	ما العواقب، والحجم، والحساسية، والأتمتة، وإمكانية الاعتراض، ومخاطر إمكانية الوصول؟
قاعدة القرار	كيف ستدمج الدرجات وتوزن وتطبق العتبات وتراجع وتُتجاوز عند الضرورة؟
الأدوار والمساءلة	من يملك القرار، والعلم، والبيانات، والخصوصية، وإمكانية الوصول، والتكنولوجيا، والتشغيل، والتدقيق؟
التحقق والمراقبة	ما الأدلة السابقة للإطلاق، والتجريب، والدراسة المحلية، وتحليل المجموعات الفرعية، والمؤشرات المستمرة المطلوبة؟
عرض المرشح	ماذا سيُقال للمشاركين؟ ما الدعم المتاح؟ وما التغذية الراجعة أو المراجعة التي ستقدم؟
دورة الحياة والخروج	كيف ستدار الإصدارات والتغييرات والحوادث والاحتفاظ بالبيانات وإنهاء علاقة المورد وإيقاف الخدمة؟

عرّف النجاح كبطاقة أداء متوازنة

البعد	مؤشرات توضيحية
جودة القرار	العلاقات بالمعايير، وجودة التعيين، والنجاح في التدريب، ونتائج الترقية، واتساق القرار، والصدق الإضافي.
العدالة والوصول	الإكمال حسب المجموعة، وتجربة التيسيرات، ومعدلات الاختيار، والتنبيؤ حسب المجموعات، وDIF، والاعتراضات، والمعالجات.
تجربة المرشح	الوضوح، والإحساس بالصلة، وإمكانية الوصول، والإكمال، والانسحاب، واتصالات الدعم، وفائدة التغذية الراجعة.
الأداء التشغيلي	زمن الإكمال، ومعدل الفشل، واتفاق المقيمين، وزمن الإنجاز، والتكلفة، والالتزام بالعملية.
الحوكمة	التوثيق، والإصدارات المعتمدة، وحادثة التدريب، وسجلات الوصول، والحوادث، ونتائج التدقيق، وتواريخ المراجعة.
القيمة التجارية	خفض تكلفة الشواغر، وتحسين الأداء، وتقليل الأخطاء أو التسرب، واتساق التنقل، وثقة المديرين.

حقوق القرار أهم من مصفوفة RACI عامة

القرار	المالك المسؤول	المدخل المستقل المطلوب
اعتماد البنية والاستخدام	مالك عملية الأعمال أو الموارد البشرية.	أخصائي تقييم، وخبراء الوظيفة، ومراجعة الشمول.
اعتماد الأداة أو المورد	مالك حوكمة التقييم.	مراجعة سيكمترية، وخصوصية، وأمن، وإمكانية وصول، ومشتريات، وقانون بحسب مستوى المخاطر.
اعتماد قواعد التسجيل والحدود	قائد علم التقييم.	مالك العمل ومراجعة عدالة أو قانون.
اعتماد الإطلاق الإنتاجي	مسؤول مخاطر تنفيذي أو مفوض.	دليل على جاهزية التدريب والضوابط والتدريب والمراقبة.
اعتماد تغيير جوهري	سلطة التغيير.	إعادة تحقق وتحليل أثر متناسب مع التغيير.
التعليق أو الإيقاف	مالك المخاطر.	مدخلات التشغيل والقانون والخصوصية والمورد والتواصل.

جرب قبل الإنتاج

ينبغي للتجريب أن يختبر أكثر من البرمجيات. يجب أن يقيّم التعليمات، ومواد التدريب، والوقت، وإمكانية الوصول، والتسجيل، وسلوك المقيمين، والدعم، وتدفعات البيانات، والتقارير، وقواعد القرار، وتجربة المرشح. وعندما يكون ذلك ممكناً، استخدم الدرجات في وضع «ظل» قبل أن تؤثر في القرارات. يسمح ذلك بمقارنتها بالأدلة الحالية واكتشاف المشكلات التقنية أو الفروق بين المجموعات من دون الإضرار بالمشاركين.

ممارسة insyt2: خطط للدليل قبل أن ترى النتيجة

نكتب أسئلة التحقق والمراقبة قبل فحص نتائج التجريب، مما يقلل إغراء إعادة صياغة القصة حول النتائج التي تبدو الأقوى.

الحد الأدنى لبوابة الإطلاق

- الغرض، والفئة، وتعريفات البنى، ودور التقييم في القرار موثقة. • أدلة تحليل الوظيفة واختيار الأساليب متاحة وتمت مراجعتها. • تم تقييم الوثائق الفنية والمعايير والتسجيل والثبات والصدق والعدالة بصورة مستقلة. • تم اختبار إمكانية الوصول وإجراءات التيسير من البداية إلى النهاية. • تم اعتماد تواصل المرشح، والدعم، وإعادة الاختبار، والحوادث، والمراجعة، ومسارات الاعتراض. • تم رسم تدفقات البيانات، وإشعار الخصوصية، والاحتفاظ، والوصول، والأمن، والتزامات المورد. • تلقى المستخدمون والمقيمون وصانعو القرار التدريب وتم التحقق من كفاءتهم. • تحققت معايير التجريب أو تم قبول الانحرافات صراحة من مالك المخاطر المسؤول. • مؤشرات المراقبة، والمسؤولون، والعتبات، وتواتر المراجعة، ومحفزات التعليق كلها مفعلة.

10

الفصل

اختيار الاختبارات وتقييم المؤردين

اشترِ الدليل والمساءلة، لا مجرد منصة ولوحة معلومات.

المشتري يظل مسؤولاً

في insyt2 نؤمن بأن...

لن نقوم insyt2 أبداً بـ...

نقل المسألة إلى المورد لمجرد أن المنتج واسع الاستخدام أو قوي التسويق.

التعاقد الخارجي على التنفيذ لا يعني التعاقد الخارجي على مسؤولية صاحب العمل عن القرار. قد يملك المورد الأداة أو المنصة أو النموذج أو بنك المفردات، لكن العميل هو من يختار الاستخدام والفئة وقاعدة القرار والدمج والتواصل والاستجابة للأدلة. يجب أن تدعم العقود التدقيق، والمراقبة، والاستجابة للحوادث، وحقوق البيانات، وضبط الإصدارات، والخروج - لا أن تمنعها.

حزمة أدلة العناية الواجبة

المجال	الأدلة المطلوب طلبها
الغرض والبنية	الاستخدامات المقصودة والمستبعدة، وتعريفات البنى، والنظرية، والربط بالوظيفة، والفئات المستهدفة، وتوصيات القرار.
المحتوى والتطوير	المخطط، وتطوير المفردات أو التمارين، ومؤهلات الخبراء، وتصميم التجريب، وتحليلات المفردات، والترجمة، والتكييف.
التسجيل والنماذج	قواعد التسجيل، وبناء المقاييس، ونموذج المقيمين، وخصائص الخوارزمية، وقواعد البيانات المفقودة، وفترات الثقة، وسجل الإصدارات.
الثبات والدقة	أدلة لكل درجة مبلّغ عنها، وتفاصيل العينة، والدقة الشرطية، واتفاق المقيمين أو تكافؤ النماذج، واتساق التصنيف.
الصدق	أدلة المحتوى، وعمليات الاستجابة، والبنية، والمعيان، والبناء، والعواقب؛ وعدم اليقين، والتحقق المتقاطع، والقيود.
المعايير والمستويات	مصدر المعايير، والتواريخ، وأخذ العينات، والتركيب السكانية، والأدوار، والدول، والمتقدمون مقابل الموظفين، ومنطق حدود القرار.
العدالة وإمكانية الوصول	DIF أو التكافؤ، والتنبؤ والنتائج حسب المجموعات، ومراقبة الأثر السلبي، واختبار إمكانية الوصول، والتوافق مع التيسيرات.
الضمان الرقمي والذكاء الاصطناعي	مصدر البيانات، ومنطق الخصائص، وضوابط التسرب، والأداء حسب المجموعات، وقابلية التفسير، والانحراف، والمراجعة البشرية، وحوكمة التحديث.
الأمن والخصوصية	ضوابط الأمن، والتشفير، ونزاهة الاختبار، والوصول، والتسجيل، ومواقع البيانات، والمعالجون الفرعيين، والاحتفاظ، والحذف، وإجراءات الاختراق.
التشغيل والدعم	مستويات الخدمة، والتوافر، ودعم الأجهزة، ومعالجة الحوادث، ودعم المشاركين، وتدريب المقيمين، والتكاملات، والتقارير.
الحوكمة والعقد	حقوق التدقيق، وإشعار التغيير، والوصول لأدلة التحقق، وحدود الملكية الفكرية، والمسؤولية، والخروج، وقابلية نقل البيانات.

اطرح أسئلة تجبر على التحديد

- ما تفسيرات الدرجات والاستخدامات الوظيفية المدعومة تحديداً، وما الاستخدامات المحظورة؟
- اعرض أحجام العينات، والتواريخ، والفئات، والوظائف، والدول، واللغات، والمعايير، والتقدير، وعدم اليقين، والتحقق المتقاطع لكل ادعاء.
- ما الخصائص أو السلوكيات التي تؤثر في الدرجة، ولماذا ترتبط بالوظيفة؟
- ماذا يحدث للثبات، والصدق، وأداء المجموعات الفرعية، وعبء المشارك عند تطبيق حد الوقت ونمط التقديم المقترحين؟
- كيف تتأثر الدرجات بالتيسيرات، والتقنيات المساعدة، والأجهزة المحمولة، والانقطاعات، وإعادة الاختبار، واستخدام الذكاء الاصطناعي التوليدي؟
- ما التغييرات التي يستطيع المورد إجراؤها من دون موافقة العميل؟ وكيف يتم تعريف الإصدارات والنماذج والمعايير والتقارير؟
- ما الأدلة التي يستطيع العميل الاحتفاظ بها أو إعادة إنتاجها إذا انتهى العقد؟
- ما المراجعات المستقلة للاختبار، أو عمليات التدقيق، أو الدراسات المحكمة المتاحة؟ وما القيود التي حددتها؟

بطاقة تقييم موزونة

حدد أوزان المعايير قبل عروض الموردين حتى لا يطفئ جمال الواجهة على الأولويات العلمية والتشغيلية. في الاختيار عالي المخاطر، ينبغي عادة إعطاء وزن أكبر للصلة بالوظيفة، والصدق، والعدالة، وإمكانية الوصول، والحوكمة، والخصوصية، والمراقبة مقارنة بحدثة الواجهة. استخدم بوابات حد أدنى إلى جانب المجموع الموزون؛ فلا يمكن لتعثر خطير في الخصوصية أو الصدق أن تعوّضه تجربة مرشح جميلة.

المعيار	وزن توضيحي	بوابة إلزامية؟
الصلة بالوظيفة ووضوح البنية	15%	نعم
أدلة الصدق للاستخدام المقصود	15%	نعم
الثبات ودقة الدرجة	10%	نعم
العدالة وأدلة المجموعات الفرعية وإمكانية الوصول	15%	نعم
شفافية التسجيل وحوكمة الذكاء الاصطناعي	10%	نعم عند التسجيل المؤتمت
حماية البيانات والأمن ونزاهة الاختبار	10%	نعم
تجربة المرشح والإدارة	8%	لا
الملاءمة التشغيلية والتكاملات	7%	لا
المراقبة والتقارير وحقوق التدقيق	5%	نعم
القيمة التجارية والتكلفة الإجمالية	5%	لا

إشارات تحذير في التسويق

الادعاء أو السلوك	لماذا يثير القلق؟
«مثبت علمياً» من دون دليل فني	لا يمكن تقييم الادعاء أو إعادة إنتاجه.
«خالٍ من التحيز» أو «يلغي التحيز البشري»	لا يوجد تقييم خالٍ من الافتراضات أو قيود البيانات أو مخاطر التنفيذ.
رقم صدق واحد لكل دور وكل دولة	الصدق يعتمد على البنى والمعايير والفئة والسياق والاستخدام.
الاستناد إلى سرية الخوارزمية لرفض الإفصاح الفني	قد تحمي الأسرار التجارية الشفرة، لكنها لا تستبدل الأدلة والمساءلة.
عرض مخططات عامة للتكافؤ الديموغرافي فقط	تكافؤ النتائج وحده لا يثبت الصلة بالوظيفة أو تكافؤ القياس أو التنبؤ أو قانونية الاستخدام.
عدم وجود سجل إصدارات أو عملية لإشعار التغيير	قد يتغير النظام الذي يجري التقييم به من دون إعادة تحقق أو علم العميل.
تسميات ذات طابع سريري في منتج مهني	قد تشجع التشخيص أو الوصم أو قرارات تتجاوز الدليل ونطاق الاستخدام.
وصف المعايير فقط بأنها «كبيرة» أو «عالمية»	لا يمكن الحكم على الصلة، أو أخذ العينات، أو التواريخ، أو التمثيل.
الضغط لتجاوز التجريب لأن عملاء آخرين يستخدمون الأداة	استخدام مؤسسة أخرى ليس دليلاً محلياً ولا ضابط تنفيذ.

مقارنة مجهولة الهوية بين موردين باستخدام إطار أدلة شفاف مشروع قوي للمحافظة المهنية، لأنه يثبت الثقافة السيكومترية والحكم التجاري في الوقت نفسه.

**منظور insyt2
للمسيرة
المهنية: اجعل
العناية الواجبة
مرئية**

الفصل

11

تطوير التقييم

ابن الأداة من البنية والمخطط، مروراً بالتجريب والتحليل والتحقق، وصولاً إلى إصدار مضبوط ومحكوم.

دورة حياة التطوير

في insyt2 نؤمن بأن...

لن نقوم insyt2 أبداً بـ...

التعامل مع كتابة المفردات كأنها كتابة تسويقية، أو إطلاق منتج قبل توثيق التسجيل والأدلة والحوكمة.

1. عرّف الغرض، والفئة، والمخاطر، والبنية، والتفسيرات المقصودة.
2. أكمل تحليل الوظيفة أو المجال وأنشئ نموذجاً للأدلة.
3. أنشئ مخططاً يحدد المحتوى، والصعوبة، والصيغة، والأسلوب، وتغطية التسجيل.
4. اكتب مواصفات المفردات أو التمارين قبل كتابة المحتوى نفسه.
5. درّب الكتاب والخبراء، وأنشئ مواد أكثر مما يحتاجه النموذج النهائي.
6. نفذ مراجعات فنية، ومحتوى، ولغة، وشمول، وإمكانية وصول، وأمن، وقانون.
7. نفذ مقابلات معرفية، واختبارات قابلية استخدام، وتجارب صغيرة.
8. نفذ تجربة تجريبية في ظروف الاستخدام المقصودة مع عينة مناسبة وخطة للمعيار.
9. حلل المفردات، والمقاييس، وعمليات الاستجابة، والثبات، والبنية، والعدالة، والبيانات التشغيلية.
10. راجع وأعد التجريب عند الحاجة بدلاً من تحسين الأداة على عينة واحدة فقط.
11. تحقق من تفسيرات الدرجات والقرارات والعواقب المقصودة.
12. ضع المعايير أو المستويات، وقواعد التقارير، والتدريب، والأدلة، والمراقبة، وضبط الإصدارات.
13. أطلق عبر عملية حوكمة، واستمر في تقييم الأداء.

مواصفات المفردات والتمارين

المواصفة وصفة تربط كل مفردة بالمخطط. فهي تحدد جانب البنية، وسياق العمل، والمثير، ومهمة الاستجابة، والصعوبة، والعملية المعرفية، والمفتاح أو منطق التسجيل، والمحتوى المحظور، ومتطلبات إمكانية الوصول، ومعايير المراجعة. تحسن المواصفات الاتساق بين الكتاب وتسمح بإنتاج محتوى متوازن بصورة مضبوطة. وتزداد أهميتها عندما يستخدم الذكاء الاصطناعي التوليدي لصياغة مسودات: يمكن للذكاء الاصطناعي تسريع الإنتاج، لكن الخبراء المدربين يجب أن يسيطروا على المواصفة، ويتحققوا من الدقة، ويحموا المحتوى السري، ويقيموا الناتج تجريبياً.

حقل الموصافة	مثال توضيحي لمحتوى اختبار حكم موقعي
جانب البنية	الحكم التدريبي: التشخيص قبل وصف الحل.
الموقف	انخفضت جودة أداء موظف خبير بعد تغيير في العملية بينما يواجه الفريق ضغطاً.
العملية المستهدفة	تحديد الأسباب، والبحث عن أدلة، وإشراك الموظف، والموازنة بين المخاطر الفورية والتعلم.
صيغة الاستجابة	تقييم أربعة تصرفات، أو اختيار الأكثر والأقل فاعلية.
أساس المفتاح	إجماع خبراء مع مراجعة تجريبية؛ البدائل تمثل مستويات مختلفة من جودة الحكم.
ضبط الصعوبة	إدراج أولويات متنافسة ومعلومات غير مكتملة من دون الاعتماد على سياسة غامضة.
ضوابط العدالة	تجنب التعبيرات المحلية الخاصة بثقافة معينة، وإشارات المكانة غير الضرورية، وافتراسات الرعاية الأسرية، والتحميل اللغوي غير ذي الصلة.
ضوابط الأمن	إنشاء بدائل للموقف وتصنيف لانكشاف المفردات، وتقييد الوصول إلى منطق التسجيل.

كتابة مفردات القدرة ذات الاستجابة المختارة

- استخدم محتوى يمثل المخطط، لا ما يكون أسهل في الكتابة.
- اجعل السؤال واضحاً ومتكاملاً، وأزل التعقيد غير المرتبط إلا إذا كان جزءاً من البنية.
- تأكد من وجود إجابة واحدة يمكن الدفاع عنها وفق الافتراضات المعلنة.
- أنشئ بدائل مشتتة من أخطاء أو مفاهيم خاطئة معقولة، لا من خيارات هزلية أو خاطئة بوضوح.
- تجنب الإشارات الناتجة عن القواعد النحوية، أو طول الخيار، أو موضعه، أو تكرار الصياغة، أو الدقة الرقمية.
- راجع مستوى القراءة، والافتراضات الثقافية، والمرئيات، والوحدات، وعرض الأجهزة، وسلوك التقنيات المساعدة.
- جرّب زمن المفردة وعمليات الاستجابة، ولا تستنتج الصعوبة من حكم الكاتب وحده.

كتابة مفردات الشخصية والاتجاهات

- اكتب عبارات سلوكية محددة ضمن إطار زمني وسياق مناسبين.
- تجنب العبارات المزدوجة، والمطلقات، والمصطلحات الغامضة، والوعظ الأخلاقي، والعبارات المرغوبة اجتماعياً بشكل شامل.
- وازن جوانب البنية مفاهيمياً؛ ولا تعتمد على الصياغة العكسية باعتبارها الضابط الرئيسي لأسلوب الاستجابة.
- تحقق من أن خيارات الاستجابة مرتبة ومتمايزة وقابلة للاستخدام عبر الثقافات والمستويات الوظيفية.
- استخدم المقابلات المعرفية لفهم كيف يفسر المستجيبون المفردة، لا ما إذا كانوا يحبونها.
- احتفظ بالمفردات بالاستناد إلى الأدلة السيكمترية وتغطية المجال معاً.

تسجيل التمارين والاستجابات المنشأة

ينبغي لأدلة التسجيل أن تحدد الأدلة القابلة للملاحظة، ومستويات الأداء، والأخطاء الحرجة، وقواعد دمج المؤشرات. المقاييس ذات المؤشرات السلوكية مفيدة عندما تصف المؤشرات فروقاً محددة في الجودة بدلاً من صفات عامة مثل «ممتاز» أو «ضعيف». دزب المقيمين باستخدام أمثلة معيارية، وممارسة مستقلة، ومناقشة الأخطاء، وفحص للكفاءة. راقب الشدة، والاتساق، والهالة، وتقييد المدى، والانحراف مع الوقت. ويجب تقييم التسجيل الآلي مقابل أحكام بشرية عالية الجودة ومعايير وظيفية، لا مجرد تدريبه على إعادة إنتاج تقييمات تاريخية قد تحتوي على تحيز.

الدليل الفني جزء من المنتج

ينبغي لأي تقييم مهني أن يمتلك وثائق كافية لتمكين مستخدم مؤهل أو مراجع من فهم الغرض، والمحتوى، والتطوير، والإدارة، والتسجيل، والمعايير، والثبات، والصدق، والعدالة، والقيود، والاستخدامات المناسبة وغير المناسبة، ومتطلبات التدريب، والأمن، وسجل الإصدارات. يجب أن تميز الوثائق بين الدليل والاستنتاج، وأن تفصح عن النتائج غير المواتية أو غير الحاسمة. قائمة مراجع طويلة ليست كافية؛ يجب أن يوضح الدليل كيف تم تقييم الأداة الفعلية والاستخدام الفعلي.

ضبط التغيير لدى insyt2: قد يعني الإصدار الجديد الحاجة إلى أدلة جديدة

التغييرات في المفردات، أو الوقت، أو خيارات الاستجابة، أو التسجيل، أو المعايير، أو المنصة، أو الترجمة، أو التقارير، أو قواعد القرار قد تغير معنى الدرجات، ولذلك يجب مراجعتها وفقاً لذلك.

الفصل

12

الإدارة، وإمكانية الوصول، وتجربة المرشح

تحمي الشروط المعيارية، والتواصل المحترم، والتكنولوجيا القابلة للاستخدام، والمعالجات الموثقة معنى الدرجة.

التوحيد في الممارسة

لن تقوم insyt2 أبداً بـ...

الخلط بين المعاملة المتطابقة والعدالة، أو إجراء تغييرات ارتجالية تغير ما يتم قياسه.

في insyt2 نؤمن بأن...

التوحيد يعني ضبط الشروط ذات الصلة بحيث تصبح الدرجات قابلة للمقارنة. ولا يعني تجاهل السياق أو إجبار كل مشارك على المرور بعملية لا يستطيع الوصول إليها. يجب أن يحدد البروتوكول التعليمات، والتدريب، والوقت، وفترات الراحة، والمواد والأدوات المسموحة، والتحقق من الهوية أو الأهلية، والبيئة، ومتطلبات الأجهزة، والدعم، والتيسيرات، والانقطاعات، وإعادة الاختبار، والسرية، وتصعيد الحوادث. وأي مساحة تقدير تُترك للمشرفين يجب أن تكون موصوفة ومشمولة بالتدريب.

ما الذي ينبغي أن يعرفه المشاركون؟

- غرض التقييم وكيف سيؤثر في العملية.
- طبيعة الأنشطة، والمدة التقريبية، وقواعد الوقت، والاستعداد أو الأدوات المسموحة.
- ما البيانات التي تُجمع، وكيف تُسجل أو تُراجع، ومن يتلقى النتائج، ومدة الاحتفاظ، والحقوق ذات الصلة.
- المتطلبات التقنية وطريقة لاختبار الجهاز أو طلب بديل.
- كيفية طلب تيسير من دون إفصاح غير ضروري.
- ما الذي يحدث بعد حادث تقني، أو إكمال متأخر، أو مخالفة مشتبّه بها، أو طلب مراجعة.
- ما إذا كانت تغذية راجعة ستقدم وما الذي ستتضمنه.
- قناة اتصال للأسئلة لا تخل بأمن المفردات.

الاستعداد مقابل التدريب على الإجابة

الاستعداد الجيد يزيد الإلمام بالصيغة ويقلل القلق غير ذي الصلة أو الحرمان الرقمي. وفر مفردات تدريبية ممثلة، وإرشادات للتنقل، وتوقعات الوقت، ومعلومات مناسبة عن البنية. تجنب كشف المفردات الحية أو مفاتيح الإجابة أو التدريب الذي يحول التقييم إلى اختبار للقدرة على الوصول إلى مصادر مميزة. راقب ما إذا كان الإعداد التجاري يغير الدرجات أو أنماط المجموعات الفرعية، وفكر في بنوك مفردات أكبر، أو نماذج بديلة، أو عينات عمل أصيلة، أو مراحل تحقق عندما يكون الانكشاف جوهرياً.

أنماط الإشراف: حضوري وعن بُعد ومن دون مراقبة

النمط	المزايا	المخاطر والضوابط
إشراف حضوري	بيئة مضبوطة، والتحقق من الهوية، والدعم التقني، وملاحظة الحوادث.	عبء السفر والوصول، واختلاف الغرف والمشرفين؛ وحد المواقع والتعديلات.
إشراف مباشر عن بُعد	مرونة جغرافية مع دعم فوري.	الخصوصية، والاتصال، والمراقبة، وبيئة المنزل؛ قلل التدخل ووفر بدائل.
تسجيل ثم مراجعة	مراجعة قابلة للتوسع للجلسات التي ترفع إشارات.	إشارات خاطئة، وخصوصية، وتحيز، وقرارات غامضة؛ استخدم قواعد واضحة ومراجعة بشرية.
اختبار إلكتروني غير مراقب	مريح، وقابل للتوسع، وأقل عبأ.	مساعدة خارجية، وعدم يقين في الهوية، واختلاف الأجهزة، والانقطاعات؛ صمم تحقّقاً وسياسة للحوادث.
اختبار تحقق	يؤكد درجة سابقة تحت شروط مضبوطة.	آثار التدريب والعبء؛ استخدم نماذج بديلة وقواعد محددة مسبقاً للفروق.

إدارة الحوادث

الحادث	الإجراء الفوري	القرار والسجل
انقطاع المنصة	احتفظ ببيانات الاستجابة والطوابع الزمنية والسجلات وقدم الدعم.	استئناف أو إعادة جدولة أو إلغاء وفق قواعد شدة محددة مسبقاً.
انقطاع بيئي	أوقف الاختبار إن أمكن وسجل رواية المشارك من دون لوم.	قيّم أثره على البنية وقدم علاجاً عندما يكون جوهرياً.
اشتباه في مساعدة خارجية	لا تنهم تلقائياً؛ احفظ الأدلة وقيد الوصول.	مراجعة مستقلة، وربما اختبار تحقق، وتوثيق مستوى الثقة ومخاطر الإيجابيات الكاذبة.
فشل التيسير	أوقف أو عدل بموافقة المشارك وأشرك مسؤول إمكانية الوصول.	وفر بديلاً صالحاً أو إعادة اختبار، وحقق في فشل العملية.
خطأ في مفردة أو مفتاح	علّق المحتوى المتأثر وحدد المشاركين المعرضين.	أعد التسجيل أو التقييم، وأخطر أصحاب العلاقة، ووثق السبب الجذري والإجراء التصحيحي.
حادث بيانات أو أمن	فعل استجابة الأمن والخصوصية واحتو الوصول.	الإخطار القانوني، ومعالجة وضع المشارك، ومراجعة النزاهة، وإعادة تشغيل مضبوطة حسب الحاجة.

تجربة المرشح ليست تجميلاً

الإحساس بالصلة، والشفافية، وفرصة الأداء، والاحترام، والاتساق تؤثر في الثقة، والانسحاب، والسمعة، وأحياناً الأداء. الواجهة الجميلة لا تعوض المحتوى غير ذي الصلة أو المراقبة الغامضة، لكن حتى الاختبار القابل للدفاع قد يفشل إذا كانت التعليمات مربكة، أو العرض على الهاتف سيئاً، أو الدعم غير متاح، أو التغذية الراجعة مهينة. ضَمّن بحث تجربة المرشح في التجريب وراقبها حسب المجموعات ذات الصلة.

نصمم رحلة المشارك بالتوازي مع المنهج السيكومتري. الدعوات، والاستعداد، والوصول، والدعم، والتيسيرات، والإكمال، والتغذية الراجعة، والمراجعة كلها تؤثر في الجودة.	معيار insyt2 للمشاركة: صمّم التجربة مع القياس
---	---

الفصل

13

التسجيل، والتفسير، والتقارير، والتغذية الراجعة

حوّل الأدلة إلى استنتاجات مفهومة ومتناسبة من دون تحويل الدرجات إلى هويات ثابتة.

ضبط جودة التسجيل

في insyt2 نؤمن بأن...

لن نقوم insyt2 أبداً بـ...
السماح للنصوص المؤتمتة بأن تبالغ في الدقة أو اليقين أو السببية.

قد تنشأ أخطاء التسجيل من مفاتيح غير صحيحة، أو قواعد البيانات المفقودة، أو عدم تطابق النموذج، أو عكس غير صحيح للمفردات، أو اختيار معيار غير مناسب، أو التحويلات، أو انحراف المقيمين، أو السجلات المكررة، أو الدمج، أو عيوب البرمجيات. تقلل الأتمتة بعض الأخطاء اليدوية لكنها تزيد أهمية المواصفات، والاختبار، وضبط الإصدارات، والمطابقة المستقلة. قبل الإطلاق، تحقق من تحويل الدرجات الخام إلى الدرجات المقيسة باستخدام حالات معروفة، وقيم حدية، وأنماط بيانات مفقودة، ونماذج بديلة، وتيسيرات، وتقارير.

الضابط	المثال
المواصفة	وثيقة تسجيل معتمدة تحدد كل رمز استجابة، ومفتاح مفردة، ووزن، وتحويل، وقاعدة بيانات مفقودة، وتسمية تقرير.
اختبار الوحدات	سجلات إدخال معروفة تنتج مخرجات متوقعة على مستوى المفردة والمقياس والمركب والتصنيف.
إعادة إنتاج مستقلة	تطبيق ثانٍ أو محلل ثانٍ يعيد إنتاج عينة من الدرجات من البيانات الخام.
مطابقة الإنتاج	فحص الأعداد وتوزيعات الدرجات بعد كل دفعة أو إصدار أو تغيير تكامل.
مراجعة القيم الشاذة	أنماط غير متوقعة، أو بيانات مفقودة، أو أزمنة شديدة، أو تحولات جماعية تطلق تحقيقاً.
ضبط الإصدار	لا تولد التقارير الرسمية إلا من إصدارات التسجيل والمعايير المعتمدة، ويحمل كل تقرير رقم إصدار.

تسلسل منضبط للتفسير

1. أكد الهوية، والغرض، وإصدار الأداة، والمعيار، واللغة، وطريقة الإدارة، والتيسير، ومؤشرات صلاحية البيانات.
2. افهم تعريف كل مقياس وما الذي لا يقيسه صراحة.
3. ضع في الاعتبار الثبات، والخطأ المعياري، ونطاق الدرجة، وما إذا كان الفرق ذا معنى.
4. فسر الأنماط باستخدام الدليل الفني ونظرية البنية، لا القصص الحدسية.
5. ادمج السياق والأدلة المؤيدة، وحقق في التناقضات بدلاً من تسويتها بمتوسط.
6. اربط الاستنتاجات بالدور أو القرار أو سؤال التطوير.
7. صرّح بالقيود والتفسيرات البديلة والمعلومات التي ما زالت مطلوبة.
8. استخدم لغة متناسبة وتجنب التسميات السريرية أو الأخلاقية أو الحتمية.

9. وثق المنطق وأي حكم مهني يتجاوز المخرجات المؤتمتة.
من الدرجة إلى السرد

صياغة ضعيفة	صياغة أقوى
«الشخص حصل على 9 في الدافعية وهو قائد بالفطرة».	«تشير الاستجابات إلى تفضيل قوي للأهداف الصعبة وبذل الجهد المستمر. استكشف ما إذا كان ذلك متوازناً بالتفويض، والاستماع، وتحديد الأولويات بصورة واقعية».
«المرشح فشل في القدرة العددية».	«جاءت الدرجة دون العتبة المحددة مسبقاً في هذه المهمة العددية المحددة بزمان. راجع نطاق الثقة وسجل الحوادث، خصوصاً إذا لم تكن السرعة العددية مطلباً جوهرياً».
«انخفاض التعاطف يعني ضعف الملاءمة للفريق».	«يشير ملف التقرير الذاتي إلى أسلوب أكثر تركيزاً على المهمة وأقل تعبيراً عاطفياً. استكشف ذلك من خلال سلوك مرتبط بالوظيفة بدلاً من اعتباره دليلاً مباشراً على فاعلية العمل الجماعي».
«الخوارزمية تتنبأ بنسبة نجاح 83%».	«يضع ناتج النموذج المرشح داخل النطاق المحدد. تنطبق الاحتمالية فقط ضمن الفئة والمعيار والنموذج والمعايرة التي تم التحقق منها؛ وليست ضماناً فردياً».

التقارير الفردية

ينبغي للتقرير أن يحدد الغرض، والأداة والإصدار، وتاريخ الإدارة، والمعيار أو المستوى، وتعريفات المقاييس، ووحدات الدرجات، والدقة، ونطاق التفسير، والقيود، والجمهور المقصود. في القرارات عالية المخاطر، ميّز بين البيانات الواقعية، والتفسير المعياري، والمخرج الخوارزمي، والحكم المهني للممارس. تجنب التفاصيل الحساسة غير الضرورية وتأكد من أن المستلم يفهم أن التقرير مرتبط بوقت وسياق وغرض محددين.

حوارات التغذية الراجعة

التغذية الراجعة تدخل مهني، وليست قراءة للتقرير. ابدأ بالغرض، والسرية، وتجربة المشاركة. اشرح النموذج ومقاييس الدرجات بلغة واضحة. استكشف الأنماط بالأمثلة وادعُ إلى تقديم أدلة مخالفة. اربط الرؤى بالدور أو أهداف التطوير، وحدد نقاط القوة والجوانب التي تستحق الانتباه، واتفق على إجراءات. لا تفرض الاتفاق ولا تصف الاختلاف مع التقرير بأنه «نقص في الوعي».

مرحلة التغذية الراجعة	أسئلة مفيدة
التعاقد	ما الذي سيجعل هذه الجلسة مفيدة؟ ما الذي يمكن مشاركته ومع من؟
التوجيه	هذا ما يأخذ التقييم عينة منه، وهذه طريقة مقارنة الدرجات، وهذه أهم القيود.
الاستكشاف	أين يتوافق هذا مع خبرتك؟ وأين لا يتوافق؟ ما الأمثلة التي تدعمه أو تتحدها؟
الدمج	كيف يتفاعل النمط مع الدور والسياق والأدلة الأخرى والأهداف الحالية؟
الفعل	ما الذي ستستمر فيه أو تبدأه أو توقفه أو تختبره أو تطلب تغذية راجعة حوله؟ وما الدعم المطلوب؟
الإغلاق	ما أهم ما خرجت به؟ ما الأسئلة المتبقية؟ وما الذي سيحدث للبيانات والتقرير؟

حد نطاق insyt2: التغذية الراجعة للتقييم ليست علاجاً نفسياً

قد تكشف التغذية الراجعة ضيقاً أو مخاوف صحية أو ادعاءات. استجب بإنسانية، واتبع بروتوكولات الحماية، وأحل الأمر إلى المختصين عندما يتجاوز نطاق كفاءتك.

الفصل

14

دمج الأدلة، والتحقق، والمراقبة، وحوكمة البيانات

يحتاج نظام الاختيار إلى منطق قرار شفاف ودورة حياة مستمرة للأدلة.

الدمج مشكلة تصميم

في insyt2 نؤمن بأن...

لن نقوم insyt2 أبداً بـ...

أخذ متوسط درجات غير متجانسة لمجرد أن جدول البيانات يسمح بذلك، أو التعامل مع الحكم البشري كاستثناء نهائي بلا منهج.

يمكن للأساليب المتعددة تحسين التغطية، وتثليث الأدلة، وتقليل الاعتماد على مقياس واحد غير كامل. لكنها قد تكرر البنى، وتزيد التكلفة، وتضخم فروق المجموعات الفرعية، وتخلق إشارات متعارضة، وتخفي المساءلة. صمّم قاعدة الدمج قبل رؤية النتائج، واربطها بتحليل الوظيفة، والصدق، والدقة، والعدالة، وتكاليف الخطأ.

النموذج	كيف يعمل	المزايا	المخاطر
حواجز متعددة	ينجح المرشح في كل مرحلة قبل الانتقال.	كفؤ ويحمي المتطلبات الدنيا.	لا يمكن تعويض الخطأ المبكر؛ ترتيب المراحل يؤثر في النتائج؛ والعتبات تحتاج إلى دليل.
مركب تعويضي	يعوض الأداء الأعلى في أسلوب الأداء الأدنى في آخر.	يستخدم المعلومات كاملة ويمكن أن يحسن التغطية.	قد يسمح بضعف غير مقبول في مطلب جوهري؛ وقد تكون الأوزان غامضة.
هجين	حدود دنيا أساسية مع مركب موزون لمن يجتازها.	يوازن بين المعايير والتنبؤ الأوسع.	أكثر تعقيداً؛ مناطق المراجعة والتفاعلات تحتاج إلى قواعد واضحة.
حكم شمولي منظم	تدمج لجنة مدربة الأدلة باستخدام أسئلة ومؤشرات محددة.	يمكنه التعامل مع السياق والاستثناءات المثبتة.	معرض لعدم الاتساق والتبرير بأثر رجعي إن لم يكن محكوماً.
توصية خوارزمية	يدمج نموذج المتغيرات في درجة أو ترتيب أو توصية.	قابل للتوسع ومتسق في ظروف مستقرة.	تسرب، وغموض، وانحراف، وتحيز بالوكالة، وتحيز للأتمتة، وصعوبة الاعتراض.

وحد المقاييس قبل الدمج

الدرجات الخام من أساليب مختلفة ليست قابلة للمقارنة مباشرة. لا معنى لمتوسط 70 في اختبار معرفة، و4 على مقياس مقابلة من خمس نقاط، و62 في مقياس T. حوّل الدرجات إلى مقياس مشترك باستخدام بيانات مرجعية مناسبة، ثم طبق الأوزان. تأكد من اتساق الاتجاه، وتعريف قواعد البيانات المفقودة، وأن الدرجة الأعلى تعني دائماً ما تقصده.

مثال توضيحي لمركب = 0.40 z للقدرة + 0.35 z للمقابلة + 0.25 z لعينة العمل

الأوزان مجرد مثال. الأوزان الفعلية تحتاج إلى أدلة من الوظيفة، والصدق، والثبات، والمنفعة، والعدالة، والحوكمة.

اختيار الأوزان

- ابدأ بأهمية البنى وبالمخطوط.
- افحص الثبات وتجنب منح وزن كبير لدرجات ضيقة وغير دقيقة.
- استخدم أدلة المعيار والتحقق المتقاطع عندما يسمح حجم العينة بالوزن التجريبي.

- راعِ الصدق الإضافي والتكرار بين الأساليب.
- قيم نتائج المجموعات الفرعية وما إذا كان مزيج بديل يحقق قيمة مشابهة مع أثر سلبي أقل.
- راعِ المنفعة، وتكلفة الإدارة، وعبء المرشح، وقابلية التنفيذ.
- فضل الأوزان البسيطة والمستقرة عندما تكون الأوزان التجريبية المعقدة عرضة لفرط الملاءمة أو يصعب تفسيرها.
- اختبر حساسية النتائج لأوزان بديلة معقولة ووثق المنطق.

الحكم البشري أسلوب من أساليب التقييم

اللجنة النهائية ليست خارج النظام السيكمترية؛ فهي إجراء اختبار آخر. أعط أعضاء اللجنة الأدلة نفسها وبالصيغة نفسها، وحدد الاعتبارات المسموحة، واستخدم التقديرات الفردية المستقلة قبل المناقشة، واربط الأحكام بمتطلبات الوظيفة، وسجل الأسباب، وافصل تصحيح الوقائع عن التفضيلات، وراقب الأنماط. يمكن للمناقشة أن تحسن اكتشاف الأخطاء، لكنها قد تعيد أيضاً إدخال إشارات المكانة، والألفة الشخصية، والسياسة الداخلية.

قد تكون الاستثناءات مشروعة بعد حادث موثق، أو فشل تيسير، أو بيانات غير صحيحة، أو دليل استثنائي ذي صلة بالعمل، أو ضمن منطقة مراجعة محددة بالسياسة. لكنها تصبح خطرة عندما يستطيع المديرون تجاهل الدرجات لمصلحة مرشحين مفضلين أو ابتكار مبررات باثر رجعي. حدد من يحق له التجاوز، وعلى أي أسس، وما الدليل المطلوب، ومن يراجع الأنماط حسب المدير والدور والموقع والمجموعة.

الإشراف البشري لا يعني السماح لشخص غير مدرب بالموافقة الشكلية على الأدلة أو قلبها بشكل عشوائي. يجب أن تكون المراجعة كفؤة، ومستنيرة، وموثقة، ومحددة الحدود.

**معييار insyt2
للقرار: يجب أن
تكون المراجعة
البشرية منظمة**

تصاميم التحقق

التصميم	نقاط القوة	القيود والضوابط
دراسة معيار تنبؤية	الأقرب إلى واقع الاختيار؛ القياس يسبق النتيجة.	الوقت، والتسرب من العينة، وجودة المعيار، والتعامل الأخلاقي؛ أحم الدراسة من تلوث المعيار بمعرفة الدرجات.
دراسة معيار متزامنة	أسرع ومفيدة كبداية للدليل المحلي.	تقييد مدى الموظفين الحاليين، والخبرة، وتأثيرات البقاء؛ فسر بحذر.
دراسة موجهة للمحتوى	قوية للمعرفة الخاصة بالوظيفة، وعينات العمل، والمقابلات، والعينات الصغيرة.	تعتمد على تعريف دقيق للمجال وخبراء مؤهلين وممثلين.
دراسة البنية	تختبر بنية الدرجات وعلاقاتها بالبنى المعروفة.	لا تثبت بمفردها التنبؤ الوظيفي أو منفعة القرار.
قابلية النقل	تستخدم الأدلة المتراكمة من سياقات متشابهة بما يكفي.	يجب إثبات التشابه؛ وتظل المراقبة المحلية والعدالة ضروريتين.
تقييم تجريبي أو شبه تجريبي	يمكن أن يقدّر أثر العملية أو المنفعة بصورة سببية.	قابلية التنفيذ، والأخلاقيات، والعينة، ودقة التطبيق.

جودة المعيار مهمة. قد تعكس تقييمات المشرف ملاحظة محدودة، أو هالة، أو تحيزاً ديموغرافياً، أو سياسة، أو معرفة بدرجات التقييم. وقد تعكس المؤشرات الموضوعية المنطقة، أو الوردية، أو الموارد، أو الفريق، أو الفرصة. ابن نموذج معيار يحدد الأبعاد المهمة، وفترات القياس، والمصادر، والثبات، والتلوث، والبيانات المفقودة. استخدم مؤشرات متعددة، واحجب درجات التقييم عن المقيمين قدر الإمكان.

لوحة متابعة المراقبة

المجال	المؤشرات الأساسية	محفز توضيحي للمراجعة
الفئة والاستخدام	الحجم، والدور، والموقع، واللغة، والقناة، والإكمال، وإعادة الاختبار، والتيسيرات.	تحول جوهري عن الفئة التي تم التحقق عليها أو عن الاستخدام المقصود.
سلوك الدرجات	المتوسط، والانحراف المعياري، والتوزيع، والبيانات المفقودة، والوقت، والثبات، وانكشاف المفردات، واتفاق المقيمين.	تحول غير متوقع، أو انخفاض الدقة، أو انحراف المقيمين.
العدالة	الإكمال ومعدلات الاختيار حسب المجموعة، ونسب الأثر، وDIF أو التكافؤ، والتنبؤ حسب المجموعة.	تفاوت مستمر بلا تفسير أو رحلة غير متاحة لبعض الفئات.
الصدق والنتائج	العلاقات بالمعيار، والمعايرة، ودقة القرار، والقيمة الإضافية، وجودة التعيين.	ضعف العلاقة، أو انعكاسها، أو فشل تكرارها.
التجربة والتشغيل	تقييمات المرشحين، والانسحاب، والدعم، والحوادث، والاعتراضات، وزمن الإنجاز، وتوافر الخدمة.	ارتفاع مفاجئ في الشكاوى أو الأعطال أو الانسحاب التفاضلي.
الحوكمة	الإصدارات، والتدريب، والوصول، والاستثناءات، والتغييرات، ومشكلات التدقيق، وتواريخ المراجعة.	استخدام غير مصرح به، أو تدريب منتهي، أو استثناءات مفرطة، أو تغيير غير معتمد.

محفزات إعادة التحقق

- تغيير جوهري في الدور، أو نموذج الكفاءات، أو توقعات الأداء، أو المعيار.
- دولة أو لغة أو فئة أو مستوى أو قرار توظيف جديد.
- تغيير في المفردات، أو التمارين، أو حدود الوقت، أو التسجيل، أو خصائص النموذج، أو الأوزان، أو المعايير، أو العتبات، أو التقارير.
- الانتقال من تقديم مراقب إلى تقديم عن بُعد أو عبر الهاتف، أو إدخال أسلوب مراقبة جديد.
- أدلة على انكشاف المفردات، أو التدريب المكثف، أو التلاعب بالاستجابات، أو مساعدة الذكاء الاصطناعي التوليدي.
- تفاوت جوهري بين المجموعات، أو فشل في إمكانية الوصول، أو شكوى، أو حادث سلبي، أو تغيير قانوني.
- انخفاض الثبات، أو القدرة التنبؤية، أو المعايرة، أو جودة التجربة، أو النزاهة التشغيلية.
- استحواذ على المورد، أو ترحيل المنصة، أو إعادة تدريب النموذج، أو تغيير المعالج الفرعي، أو فقدان الأدلة الفنية.

أمن الاختبار والخصوصية

يحمي أمن الاختبار المحتوى، والنماذج، ومفاتيح التسجيل، والخوارزميات، وإجراءات الإدارة حتى تبقى الدرجات قابلة للمقارنة. أما أمن المعلومات فيحمي سرية البيانات الشخصية والتنظيمية وسلامتها وتوافرها. تتداخل الضوابط لكنها ليست متطابقة. نشر مفردة حية قد لا يكشف بيانات شخصية لكنه قد يبطل صلاحية الدرجات المستقبلية؛ بينما تخزين فيديوهات المراقبة قد يدعم التحقق من الهوية لكنه يخلق مخاطر خصوصية وأمن.

مرحلة دورة الحياة	أسئلة البيانات والحوكمة
الجمع	ما الاستجابات والأزمنة والنقرات والصوت والفيديو وبيانات الجهاز والهوية والتيسيرات والبيانات الديموغرافية والسياقية التي تُجمع؟ وهل كل عنصر ضروري؟
النقل	ما الأنظمة والدول وواجهات API والمعالجون الفرعيون والمشرفون الذين يتلقون البيانات؟ وهل النقل مشفر ومسجل؟
المعالجة	كيف تتحول البيانات الخام أو تُسجل أو تُرفع عليها إشارات أو تُدمج أو تُستخدم للاستنتاج أو تدريب النماذج؟
الاستخدام والإفصاح	من يرى الاستجابات الخام، والدرجات، والتقارير، والإشارات، أو القرارات؟ وما الذي يشارك مع المديرين أو الموردين أو الباحثين أو العملاء؟
الاحتفاظ	ما مدة الحاجة إلى كل نوع من البيانات للقرارات والاعتراضات والتحقق والالتزامات القانونية والمراقبة؟
الحذف أو إخفاء الهوية	من يطلق الحذف؟ وكيف يتم التحقق منه عبر النسخ الاحتياطية والمعالجين الفرعيين؟ ومتى تكون البيانات مجهلة فعلاً؟
الخروج	ما الذي يُعاد أو يُتلف عند انتهاء المورد أو الغرض؟ وهل يمكن إعادة إنتاج القرارات التاريخية وأدلتها؟

اختبار تقليل البيانات

1. هل عنصر البيانات ضروري للغرض المعلن؟
2. هل هو صالح وذو صلة بالعمل، لا مجرد متاح أو تنبؤي في مجموعة بيانات واحدة؟
3. هل يمكن لإشارة أقل تدخلاً أن توفر قيمة مشابهة؟
4. هل يمكن أن يكشف أو يعمل كمؤشر بديل لخاصية حساسة؟
5. هل يفهم المشارك بصورة معقولة عملية الجمع والاستخدام؟
6. هل تستطيع المؤسسة تفسير أثر هذا العنصر في الدرجة والاعتراض عليه؟
7. هل يمكن تأمينه، والاحتفاظ به، وتصحيحه، وحذفه بصورة مناسبة؟
8. هل ستشعر المؤسسة بالارتياح لو وصفت هذا الاستخدام علناً وللشخص المتأثر؟

**مبدأ insyt2
للبيانات: لا بيانات
من دون غرض**

لا نجمع بيانات إضافية «لعلها تصبح مفيدة لاحقاً». كل عنصر بيانات يحتاج إلى غرض محدد، ودليل، ومالك، وقاعدة وصول، ومدة احتفاظ، ومسار حذف.

أساليب التقييم

كيف تعمل الأساليب الرئيسية، وأين تضيف قيمة، وأين يبالغ الممارسون غالباً في استخدامها.



في هذا الجزء

15 أساليب الأداء الأقصى

16 أساليب الأداء المعتاد

17 أدوات التطوير والتنظيم

العلم • العدالة • جودة القرار • الحكم المهني

رؤية insyt2

السؤال ليس أي اختبار يبدو أكثر إبهاراً، بل أي دليل يخدم القرار، والأشخاص المتأثرين به، وما يحدث بعد ذلك بصورة أفضل.

15

الفصل

أساليب الأداء الأقصى

تقيس اختبارات القدرة والمعرفة، وعينات العمل، والمحاكاة، واختبارات الحكم الموقفي، والمقابلات، ومراكز التقييم ما يستطيع الأشخاص إظهاره.

اختبارات القدرة المعرفية والاستعداد

أسئلة للممارس	خيار التصميم
هل السرعة متطلب جوهري، أم أن الوقت يضيف أساساً قلقاً أو تبايناً بسبب الإعاقة أو اللغة؟	اختبار سرعة مقابل قوة
هل يمثل الاستدلال العام الدور بصورة أفضل أم استعداد متخصص في مجال معين؟ وأيها يضيف دليلاً مميزاً؟	عام مقابل محدد
هل يحسن سياق الوظيفة الصلة أم يدخل معرفة سابقة وألفة ثقافية غير ضرورية؟	مجرد مقابل سياقي
هل معايرة بنك المفردات، ومسارات التكيف، والأمن، وشرح التجربة للمرشح، والدقة الشرطية قوية بما يكفي؟	ثابت مقابل تكيفي
ما ضوابط الهوية، والمساعدة الخارجية، والأجهزة، والانقطاعات، والتحقق المتناسبة مع المخاطر؟	مراقب مقابل غير مراقب
هل القرار مقارنة تنافسية أم حد أدنى للكفاءة؟ وهل المجموعة المرجعية أو حد القرار قابلاً للدفاع؟	مقارنة معيارية مقابل معيار كفاءة

اختبارات المعرفة والمهارة

- تقيس اختبارات المعرفة ما يعرفه الشخص حالياً، بينما تطلب اختبارات المهارة تطبيق المعرفة. تكون أكثر قابلية للدفاع عندما يكون المحتوى ضرورياً عند الدخول ولا يمكن تعلمه بصورة معقولة بعد التعيين. قد يقيس اختبار فني يكافى معرفة غامضة خاصة بالمؤسسة الفرصة السابقة أكثر مما يقيس القدرة المستقبلية. استخدم مخطط مجال، وحدّث المحتوى مع تغير العمل، وافصل متطلبات الدخول الحقيقية عن المعرفة القابلة للتدريب.
- استخدم مهاماً وبيانات وأدوات ممثلة للعمل مع حماية المعلومات التنظيمية السرية.
 - حدد ما إذا كانت المراجع، والآلات الحاسبة، ومساعدات البرمجة، أو الذكاء الاصطناعي التوليدي مسموحة؛ وحالِ العمل الفعلي ما لم يكن الأداء دون مساعدة هو البنية المقصودة.
 - تجنب التكاليف التي تتطلب جهداً غير مدفوع غير معقول أو تنتج عملاً ذا قيمة تجارية لصاحب العمل.
 - سجّل الاستجابات المنشأة باستخدام مقاييس واضحة وتحقق من اتساق المقيمين.
 - وفر طرقاً بديلة لإظهار المتطلب عندما لا تكون واجهة أو أداة بعينها جزءاً جوهرياً من العمل.

عينات العمل والمحاكاة

تطلب عينات العمل من المشاركين أداء مهام تشبه الوظيفة إلى حد كبير. أما المحاكاة فتنشئ بيئة مضبوطة يحدد فيها المشاركون الأولويات، أو يحلّل، أو يتواصل، أو يتفاوض، أو يشرف، أو يشخص، أو يستجيب لأحداث. قد تبدو الصلة الظاهرية عالية، لكن التشابه وحده لا يثبت الصدق. فالموقف، ومتطلبات المهمة، وفرص الملاحظة، وأبعاد التسجيل، وتدريب المقيمين، وقاعدة الدمج هي ما يحدد ما يقيسه الأسلوب فعلاً.

الأسلوب	مفيد لقياس	ضوابط الجودة
سلة الأعمال أو البريد الوارد	تحديد الأولويات، والتحليل، والتفويض، والاستجابة الكتابية.	عبء عمل موحد، ووقت، ومعلومات، ومؤشرات تسجيل، وصيغ بديلة.
لعب الأدوار	التدريب، والمبيعات، والتفاوض، والصراع، والاستجابة للعميل.	ممثلون مدربون، وتعليمات موحدة، ومؤشرات سلوكية، وملاحظات متعددة.
تحليل الحالة	تعريف المشكلة، واستخدام البيانات، والحكم، والتوصية.	حالات قابلة للمقارنة، وافتراضات صريحة، وتسجيل يفصل التحليل عن جمال العرض.
العرض التقديمي	التلخيص، والتأثير، والتواصل، والاستجابة للأسئلة.	صلة بالوظيفة، ووقت إعداد موحد، وأسئلة منظمة، وفصل محتوى العرض عن أسلوب تقديمه.
تمرين جماعي	التفاعل، والاستماع، والتأثير، وحل المشكلات تعاونياً.	ظروف جماعية متكافئة، ومقيمون متعددون، وضوابط تمنع مساواة السيطرة بالقيادة.
محاكاة فنية	البرمجة، والعمليات، والتشخيص، والتصميم، أو الاستجابة للسلامة.	أدوات ممثلة، وواقعية المهمة، والأمن، وتسجيل العملية والنتيجة.

ينبغي لأدلة تسجيل التمارين أن تعرف أدلة قابلة للملاحظة بدلاً من الصفات العامة. قد يصف المؤشر القوي كيف يحدد المشارك الأولويات المتنافسة، ويتحقق من الافتراضات، ويقيم المخاطر، ويتواصل حول القرار. أما المؤشر الضعيف فيقول فقط «حكم ممتاز». استخدم أمثلة معيارية، ومعايرة المقيمين، والتقديرات المستقلة، ومراقبة الانحراف.

اختبارات الحكم الموقفى

تعرض اختبارات الحكم الموقفى مواقف ذات صلة بالوظيفة وتطلب من المستجيب اختيار استجابات أو ترتيبها أو تقييمها أو إنشائها. قد تقيس معرفة السلوك الفعال، أو الحكم، أو الميول التفاعلية، أو مزيجاً من البنى. يعتمد معنى الدرجة على عينة المواقف، وتعليمات الاستجابة، والمفتاح، وعبء القراءة أو الوسائط، وما إذا كان المطلوب هو «الأفضل»، أو «ما ستفعله غالباً»، أو «ما فعلته سابقاً».

أسلوب المفتاح	القوة	التنبيه
إجماع خبراء الموضوع	يستند إلى حكم ذوي الخبرة وأدلة المحتوى.	قد يعيد الخبراء إنتاج الثقافة المحلية أو السياسة أو ممارسة شائعة لكنها غير فعالة.
مفتاح تجريبي	توزن الاستجابات وفق علاقتها بمعيار.	يحتاج إلى عينات كبيرة ونظيفة وضوابط ضد فرط الملاءمة وتحيز المعيار وعدم الاستقرار.
مفتاح نظري	ترتبط البدائل ببنية أو نموذج عملية محدد.	يجب أن تفسر النظرية الفروق ذات المعنى بين البدائل وأن تُختبر تجريبياً.
هجين	يجمع النظرية ومراجعة الخبراء والأدلة المعرفية والبيانات.	يجب توثيق كيفية حل التعارض بين مصادر الأدلة.

البدائل الجيدة في اختبار الحكم الموقفى كلها معقولة وتمثل فروقاً ذات معنى. تجنب إجابة أخلاقية واضحة محاطة بدائل متهورة. ضع أولويات متنافسة من دون تحويل المهمة إلى لغز. للاستخدام العالمي، كيف الافتراضات الاجتماعية والسياسات وإشارات المكانة وأعراف التواصل بدلاً من ترجمة الكلمات ميكانيكياً.

المقابلات المنظمة

تستخدم المقابلة المنظمة أسئلة مرتبطة بالوظيفة، وأسئلة متابعة موحدة، ومحاورين مدربين، ومؤشرات أدلة واضحة، ومقاييس تقدير ذات مؤشرات، وقاعدة دمج متسقة. تطلب الأسئلة السلوكية أمثلة سابقة؛ وتسال الأسئلة الموقفية عما سيفعله المرشح؛ ويمكن إدراج أسئلة فنية ودافعية عندما تكون محددة ومسجلة بوضوح. التنظيم متصل وليس ثنائياً، لكن كتابة الأسئلة مسبقاً وحدها لا تجعل المقابلة منظمة.

ثمانية ضوابط للمقابلات المنظمة

1. اشتق الأبعاد والأسئلة من تحليل الوظيفة.
2. وفر فرصاً متعددة لإظهار كل متطلب حاسم.
3. حدد أسئلة المتابعة المسموح بها بحيث توضح ولا تُلَقِّن.
4. أنشئ مؤشرات سلوكية، ومراسي للتسجيل، وإجابات معيارية.
5. درّب المحاورين على تدوين ملاحظات قائمة على الأدلة، والتقييم المستقل، وتجنب الأسئلة المتطفلة أو المرتبطة بخصائص محمية.
6. سجّل كل متطلب قبل تكوين انطباع عام.
7. ادمج الدرجات وفق القاعدة المحددة مسبقاً ووثق أي استثناء.
8. راقب اتفاق المقيمين، وتوزيعات الدرجات، ونتائج المجموعات الفرعية، وتجربة المتقدم، والعلاقات بالمعيار.

مراكز التقييم

مركز التقييم عملية موحدة وليس مكاناً فعلياً. يكمل المشاركون عدة تمارين مرتبطة بالوظيفة، ويراقب السلوك عدة مقيمين مدربين، وتُدمج الأدلة مقابل متطلبات محددة. يوم يتضمن عدة مقابلات ليس بالضرورة مركز تقييم. تعتمد الجودة على مصفوفة التمارين، والتغطية السلوكية، وكفاءة المقيمين، والتوحيد، والتسجيل، والدمج، والمراقبة.

تحتاج مراكز التقييم إلى موارد كبيرة، وينبغي استخدامها عندما يبرر اتساع الأدلة وعواقب القرار التكلفة. قد تكون تأثيرات التمرين كبيرة، لذلك قد لا تكون درجة البُعد ثابتة بالقدر نفسه عبر كل تمرين. أنشئ أكثر من ملاحظة للمتطلبات المهمة ولا تبلغ إلا بمستوى الدقة الذي يدعمه الدليل.

مقارنة أساليب الأداء الأقصى

الأسلوب	القوة الأساسية	مكفل مفيد	الخطر التصميمي الأساسي
اختبار قدرة أو استعداد	دليل كفو على الاستدلال، أو التعلم، أو استعداد محدد.	عينة عمل أو مقابلة منظمة.	سرعة أو لغة أو تعليم أو متطلب وصول غير ضروري.
اختبار معرفة	تغطية مباشرة لمعرفة دخول محددة.	إظهار مهارة أو سؤال منظم.	معلومات هامشية، ومحتوى قديم، ومكافأة الفرصة السابقة.
عينة عمل	صلة سلوكية عالية ودليل أداء مرئي.	استدلال أو مقابلة منظمة لاستنتاج أوسع.	التكلفة، وضيق العينة، والتسجيل غير الموحد، والعبء المفرط على المرشح.
اختبار حكم موقفي	يأخذ عينة من الحكم عبر عدة مواقف وظيفية بكفاءة.	مقابلة أو محاكاة تلاحظ السلوك المنفذ.	بنية غامضة، وبدائل واضحة أكثر من اللازم، وافتراضات ثقافية محلية.
مقابلة منظمة	دليل مرن عن الأمثلة والدافعية والحكم.	اختبار موحد أو عينة عمل.	عدم اتساق المتابعة، والهالة، والحكم الشمولي بأثر رجعي.
مركز تقييم	أدلة واسعة متعددة الأساليب والتمارين والمقيمين.	بيانات عمل ونتائج عالية الجودة.	التعقيد، وانحراف المقيمين، والمبالغة في تفسير الأبعاد، وعدم تكافؤ الظروف.

تسلسل العملية

ضع الأساليب المبكرة حيث تقدم معلومات مفيدة بتكلفة متناسبة وحيث يمكن الدفاع عن الرفض الخاطي. قد تكون شاشة قصيرة ومبسرة مناسبة عند أحجام ضخمة، لكن الحواجز المبكرة تنتج أيضاً أكبر استبعاد تراكمي. فكر فيما إذا كانت مرحلة أولى أوسع وأقل عبئاً، تليها أدلة أعمق، ستحافظ على عدد أكبر من المرشحين القادرين. نمذج معدلات الاختيار ونتائج المجموعات الفرعية في كل مرحلة، لا النتيجة النهائية فقط.

اختيار الأسلوب لدى insyt2: استخدم أوضح دليل بأقل عبء

اختر الأسلوب الذي ينتج أوضح دليل ذي صلة بالوظيفة مع أقل عبء ومخاطر غير ضرورية - لا الأسلوب الذي يبدو الأكثر تطوراً تقنياً.

الفصل

16

أساليب الأداء المعتاد

تصف الشخصية، والقيم، والدوافع، والاهتمامات، والنزاهة، والبيانات السيرية الميول والتفضيلات أكثر مما تصف الحد الأقصى للقدرة.

تقييم الشخصية

تطلب قوائم الشخصية المهنية من الأشخاص وصف أنماط متكررة في السلوك والتفكير والشعور. تنظم نماذج السمات الواسعة مثل نموذج العوامل الخمسة كثيراً من الميول ذات الصلة بالعمل، بينما قد تستخدم الأدوات التجارية جوانب أضيق أو تسميات خاصة. الشخصية احتمالية وسياقية. قد تدعم السمة الأداء في موقف وتخلق مخاطرة في آخر. فالحزم المرتفع قد يساعد على التأثير لكنه قد يضعف الاستماع إذا لم يُنظم جيداً، وانخفاض الاجتماعية لا يعني العجز عن التعاون.

استخدم درجات الشخصية لبناء فرضيات منظمة، لا هويات ثابتة. فسر الأنماط، ومستوى الثقة، وجودة الاستجابة، والسياق، والتفاعل مع متطلبات الدور. تجنب «الملفات المثالية» المبسطة، لأنها قد تقلل التنوع وتشجع إدارة الانطباع. في الاختيار، عرّف الفرضيات المرتبطة بالوظيفة قبل رؤية الملفات وأكدها بأدلة سلوكية. في التطوير، قدم تغذية راجعة آمنة نفسياً وساعد المشاركين على تحويل الميول إلى اختيارات واستراتيجيات مقصودة.

الصيغ المعيارية والاختيار القسري

الصيغة	الخصائص	الآثار
مقياس تقدير معياري	تُقيم كل عبارة بصورة مستقلة.	يدعم درجات السمات والمقارنات بين المجموعات، لكنه قد يتأثر بأسلوب الاستجابة وتقديم الذات.
اختيار قسري إيساطي تقليدي	ترتب البدائل أو تختار، فتننتج درجات مترابطة داخل الشخص.	قد يدعم التأمل في الملف لكنه إشكالي عموماً عند مقارنة الأشخاص بالإحصاءات التقليدية.
اختيار قسري حديث متعدد الأبعاد	تستخدم المفردات والنماذج لتقدير مستويات معيارية للسمات من الاختيارات.	قد يقلل بعض آثار الرغبة الاجتماعية الواضحة، لكنه يحتاج إلى نمذجة متقدمة وأدلة على استعادة الدرجات.
اختبار شخصية تكيفي	يتغير اختبار المفردات بناء على الاستجابات.	يمكن أن يحسن الكفاءة لكنه يحتاج إلى بنوك معايرة، وأمن، وقابلية تفسير، ودراسة دقيقة لعمليات الاستجابة.

الاختيار القسري ليس مناسباً تلقائياً للاختيار ولا محصناً من التزييف. أسأل كيف بُنيت كتل المفردات، وكيف ضُبطت الرغبة الاجتماعية، وما النموذج السيكمترى المستخدم، وهل الدرجات معيارية فعلاً، وكيف قُدِّر الثبات، وما أدلة المعيار والمجموعات الفرعية المتاحة.

تقديم الذات وجودة الاستجابة

يدرك المتقدمون غالباً أن الاستجابات قد تؤثر في الاختيار وقد يقدمون أنفسهم بصورة إيجابية. هذا لا يعني أن كل ملف مرغوب اجتماعياً غير صادق. الدافعية للوظيفة، ومفهوم الذات، والثقافة، وشفافية المفردة، والمستوى الحقيقي للسمة كلها تؤثر في الاستجابات. قد تساعد مؤشرات عدم الاتساق ومقاييس الصلاحية في المراجعة، لكنها نادراً ما تثبت الخداع.

- استخدم مفردات سلوكية محددة ومرتبطة بالعمل بدلاً من عبارات أخلاقية.
- لا تتهم الشخص بالكذب استناداً إلى مؤشر واحد.
- تعامل مع إشارات أسلوب الاستجابة بوصفها دليلاً للمراجعة، لا سبباً تلقائياً للاستبعاد.
- قيم ما إذا كانت التحذيرات أو ضوابط الاستجابة تحسن الصدق من دون زيادة القلق أو الإيجابيات الكاذبة.
- ادمج التقرير الذاتي مع أساليب سلوكية منظمة عندما تكون العواقب كبيرة.

لن تعامل insyt2 أبداً مقياس أسلوب الاستجابة بوصفه «كاشف كذب»

قد تعكس مؤشرات أسلوب الاستجابة تقديم الذات، أو الثقافة، أو التفسير، أو خداع الذات. وهي تحتاج إلى تفسير متناسب لا إلى حكم أخلاقي.

القيم والدوافع والاهتمامات

تصف القيم ما يعتبره الناس مهماً؛ وتصف الدوافع النتائج أو المحركات المرغوبة؛ وتصف الاهتمامات الانجذاب إلى أنشطة وبيئات معينة. يمكن لهذه المقاييس دعم استكشاف المسيرة، والتطوير، والارتباط الوظيفي، وحوارات الدور. وهي أقل مباشرة في قياس القدرة وقد تتغير مع السياق أو المرحلة الحياتية.

الاستخدام	تفسير مسؤول
الإرشاد المهني	أنشئ خيارات وأسئلة، وادمجها مع المهارات والقيود والخبرة ومعلومات سوق العمل.
تطوير القيادة	استكشف ما ينشط السلوك أو يهدده وكيف تتفاعل الدوافع مع القيم ومتطلبات الدور والتنظيم الذاتي.
تطوير الفريق	ناقش الفروق في الأولويات وظروف العمل من دون تنميط أو فرض الإفصاح.
الاختيار	استخدم فقط عندما تكون صلة البنية، والتسجيل، والعدالة، والتنبيه مدعومة؛ وتجنب استنساخ الثقافة الحالية.
الارتباط الوظيفي	اجمع الأنماط لتحسين تصميم العمل بدلاً من لوم الأفراد على عدم المواءمة التي تسببها المنظومة.

«الملاءمة الثقافية» غالباً مبرر غامض ومحفوف بالمخاطر. فالتشابه مع الموظفين الحاليين قد يعيد إنتاج أنماط توظيف واستبعاد سابقة. المنهج الأقوى يحدد المعايير الأخلاقية أو السلوكية الأساسية، ويفحص الظروف المطلوبة للأداء، ويقدر مساهمة الشخص في الثقافة إلى جانب قدرته على التوافق معها.

اختبارات النزاهة

تسأل اختبارات النزاهة الصريحة عن الاتجاهات نحو السلوك غير النزهي أو الضار وأحياناً عن السلوك السابق. أما اختبارات النزاهة المرتبطة بالشخصية فتستنتج المخاطر من مبول أوسع مثل الضمير المهني أو الاعتمادية أو الالتزام بالقواعد. تحتاج هذه الأدوات إلى مراجعة قانونية وثقافية وخصوصية وبنائية دقيقة. تجنب المفردات المتطفلة أو الاتهامية، والادعاءات غير المدعومة بشأن «كشف السرقة» أو «كشف عدم الأمانة»، والقرارات بلا تأكيد مستقل. عزف المعيار، وراع المعدلات الأساسية والإيجابيات الكاذبة، ووفر مراجعة بشرية ذات معنى.

البيانات السيرية

تستخدم البيانات السيرية معلومات موحدة عن الخبرات والسلوكيات السابقة. يمكن بناء مفاتيحها منطقياً انطلاقاً من تحليل الوظيفة أو تجريبياً وفق علاقتها بالنتائج. تسأل البيانات السيرية الجيدة أسئلة قابلة للتحقق، وتراعي اختلاف الفرص، ولها صلة واضحة بالعمل المستقبلي. أما الضعيفة فتكافئ الوصول المتميز، أو المسميات البراقة، أو المؤشرات غير المباشرة عن نمط الحياة. استبعد المحتوى المحمي والمتطفل، وافحص اختلاف الفرص، وامنع تسرب المعيار، وفصل التسجيل الشفاف على الارتباطات التاريخية غير المفسرة.

ممارسة insyt2: طوّر فرضيات، لا هويات ثابتة

تكون أدوات الأداء المعتاد أقوى عندما تعمق فرضية منظمة حول سلوك العمل، وأضعف عندما تتحول إلى تسميات حتمية أو استبعادات آلية.

الفصل

17

أدوات التطوير والتنظيم

تنطبق المبادئ السيكمترية نفسها، لكن إخفاء الهوية، والتجميع، والتغذية الراجعة، والفعل التنظيمي تصبح عناصر مركزية.

التغذية الراجعة بزاوية 360 درجة

تجمع عملية 360 درجات تقديرات من الشخص نفسه ومن مجموعات ذات صلة مثل المدير، والزملاء، والمرؤوسين، والعملاء، أو شركاء المشروعات. والغرض المعتاد هو التطوير: تحديد الأنماط السلوكية، واختلاف وجهات النظر، ونقاط القوة، والأولويات. ويهم نموذج الكفاءات، واختيار المقيمين، وإخفاء الهوية، وتصميم المقياس، والتعليقات، والتقارير، وتيسير التغذية الراجعة بقدر ما يهم الاتساق الداخلي.

قضية التصميم	الممارسة الجيدة
الكفاءات والمفردات	استخدم سلوكاً قابلاً للملاحظة مرتبطاً بالدور والاستراتيجية؛ وتجنب السمات الغامضة، والعبارات المزدوجة، والطول المفرط.
اختيار المقيمين	اختر أشخاصاً لديهم فرص ملاحظة كافية، وراجع العينات غير المتوازنة أو المختارة لأسباب سياسية.
إخفاء الهوية	وضّح الحد الأدنى لأحجام المجموعات والاستثناءات؛ ولا توح بسرية لا يستطيع التقرير ضمانها.
تصميم المقياس	عزّف مؤشرات للتكرار أو الفاعلية، وأدرج «لم تتم ملاحظته» عندما يكون ذلك مناسباً.
الفروق بين الذات والآخرين	تعامل مع الفجوات كفرصيات تتأثر بفرص الملاحظة والتوقعات والعلاقات واستخدام المقياس، لا كدليل قاطع على «نقاط عمياء».
التعليقات	قدم إرشادات، وأزل المحتوى الضار أو الكاشف للهوية وفق سياسة شفافة، وافصل الدليل عن التنقيص.
التغذية الراجعة	استخدم ميسرين مدربين، وركز على الأنماط والتجارب العملية، واربط الرؤية بالدعم والمتابعة.
إعادة الاستخدام	لا تحوّل بحدوء تقديرات تطويرية سرية إلى أدلة اختيار أو تأديب.

استبيانات الارتباط الوظيفي وتجربة الموظف

تقيس استبيانات الموظفين التصورات والاتجاهات والنوايا والخبرات خلال فترة زمنية. يبدأ الاستبيان القوي بنظرية تشرح ما تستطيع المؤسسة التأثير فيه وما القرارات التي ستدعمها النتائج. ويميز بين النتائج مثل الارتباط، والالتزام، والرفاه، أو نية البقاء، وبين المحركات المحتملة مثل القيادة، والموارد، والعدالة، والنمو، والشمول، وتصميم العمل. ارتباط العبارات الإيجابية بعضها ببعض لا يثبت السببية.

1. عزّف البنى ونموذج الفعل قبل كتابة المفردات.
2. اختر المسح الشامل أو أسلوب العينة بما يلائم القوى العاملة والغرض.
3. صمّم إخفاء الهوية، والسرية، والوصول، والحد الأدنى لأحجام التقارير بشفافية.
4. اختر وضوح المفردات، والترجمة، وإمكانية الوصول، ومقاييس الاستجابة، والاستخدام على الهاتف، وطول الاستبيان.
5. قيم الثبات والبنية، إلى جانب تكافؤ القياس بين المجموعات والموجات الزمنية.
6. أبلغ عن عدم اليقين، ومعدلات الاستجابة، ومخاطر عدم الاستجابة، والاتجاهات، والمعايير المقارنة، وحجم الفروق عملياً.
7. تجنب ترتيب الوحدات الصغيرة على فروق غير مستقرة أو نشر تعليقات تكشف الهوية.
8. أغلق الحلقة عبر التواصل، والفعل، والقياس اللاحق، وتقييم الأثر.

يمكن للمعايير الخارجية أن توفر سياقاً، لكنها قد تختلف في القطاع، والدولة، وتوزيع الوظائف، وطريقة الاستبيان، والتوقيت، والصياغة. يكون الاتجاه الداخلي غالباً أكثر قابلية للفعل عندما يظل المقياس والفئة قابلين للمقارنة. وقد يكون الفرق الدال إحصائياً صغيراً أكثر من اللازم ليكون مهماً، بينما قد لا تظهر مشكلة محلية خطيرة في مئين خارجي. اجمع فترات الثقة، وأحجام الأثر، والتوزيعات، وأدلة المفردات، والبيانات النوعية، وسياق الأعمال.

أدوات الفريق والثقافة

يمكن لأدوات الفرق أن تنشئ لغة مشتركة للتفضيلات، والصراع، واتخاذ القرار، والتعاون. لكنها لا ينبغي أن تتحول إلى تنميط جامد أو مبرر لتوزيع العمل بناء على ملف بدلاً من المهارة والموافقة والسياق. وينبغي لاستبيانات الثقافة أن تميز بين الأعراف الوصفية، والقيم المشتركة، وسلوك القيادة، والأنظمة، والمناخ. لا تستخدم ملف الثقافة الحالية كقالب غير قابل للنقاش للموظفين المستقبليين.

الإمكانات وتحديد المواهب

الإمكانات ليست سمة واحدة. إنها توقع بشأن القدرة المستقبلية، والدافعية، والفرصة، والجاهزية لوجهة وفترة زمنية محددتين. ينبغي لنظام المواهب أن يفصل بين الأداء الحالي، والقدرة على التعلم، والطموح، والحركة، والسلوك القيادي، والعمق الفني، والخبرة، والفرص التنظيمية. تتأثر ترشيحات المديرين بالظهور والرعاية، كما تتأثر درجات التقييم بإمكانية الوصول واختيار البنى.

السؤال	استجابة تصميم أفضل
إمكانات لأي شيء؟	عرّف الأدوار الوجهة، ومستوى التعقيد، والانتقالات الحرجة، والأفق الزمني.
متى يصبح الشخص جاهزاً؟	افصل بين «جاهز الآن»، و«جاهز مع التطوير»، والإمكانات الأطول مدى.
من أين يأتي الدليل؟	اجمع جودة الأداء، والمحاكاة، وأدلة التعلم، والدافعية، والخبرة، والحكم المنظم.
من حصل على الفرصة؟	راجع الوصول إلى المهام الممتدة، والرعاية، والظهور، والحركة، والتغذية الراجعة.
ماذا يحدث بعد ذلك؟	اربط التصنيف بتطوير شفاف، ومراجعة، وحركة، ومسارات للتصحيح.
هل نجح النظام؟	تتبع التقدم، والأداء، والاحتفاظ، والتجربة، ونتائج المجموعات الفرعية بمرور الوقت.

قاعدة insyt2 لحالة الاستخدام: صلاحية الأداة للتطوير لا تعني صلاحيتها للاختيار
قد تكون الأداة مفيدة للتأمل، ومع ذلك غير مناسبة لقرارات فردية عالية المخاطر. يجب أن يطابق الدليل الاستخدام المقصود.

الحوكمة والمستقبل الرقمي

الأخلاقيات المهنية، والسياق القانوني، وضمان الذكاء الاصطناعي، ونموذج تشغيل للتقييم المسؤول.



في هذا الجزء

18 الأخلاقيات، والقانون، والمساءلة المهنية

19 الذكاء الاصطناعي، والخوارزميات، والألعاب، والفيديو، والتقييم الرقمي

20 حوكمة التقييم، ومركز التميز، والنضج

العلم • العدالة • جودة القرار • الحكم المهني

رؤية insyt2

السؤال ليس أي اختبار يبدو أكثر إبهاراً، بل أي دليل يخدم القرار، والأشخاص المتأثرين به، وما يحدث بعد ذلك بصورة أفضل.

الفصل

18

الأخلاقيات، والقانون، والمساءلة المهنية

الامتثال القانوني يختلف باختلاف الولاية القضائية؛ والممارسة الأخلاقية تتطلب أيضاً الكفاءة، والعدالة، والشفافية، والاحترام.

الأخلاقيات ليست إخلاء مسؤولية

في insyt2 نؤمن بأن...

لن نقوم insyt2 أبداً بـ...

استخدام لغة الموافقة، أو إشعار الخصوصية، أو إخلاء مسؤولية خاص بالذكاء الاصطناعي بدلاً عن تصميم تقييم مسؤول.

المبدأ الأخلاقي	أسئلة للممارسة
الإحسان وعدم الإضرار	ما القيمة التي يمكن أن يخلقها التقييم؟ وما الأضرار المحتملة؟ وكيف ستمنع أو تُعالج؟
الاحترام والاستقلالية	ما الذي يفهمه المشاركون وما الذي يملكون السيطرة عليه؟ وما معنى الاختيار عندما تؤثر المشاركة في التوظيف؟
العدالة والإنصاف	من يستطيع إظهار البنية؟ ومن يتحمل العبء أو الخطأ أو الاستبعاد أو المراقبة؟
الكفاءة	هل يملك كل ممارس المعرفة والمؤهلات والإشراف والخبرة الحديثة للمهمة؟
النزاهة	هل الادعاءات دقيقة، والقيود معلنة، والتعارضات مدارة، والأدلة غير المواتية مذكورة؟
السرية والخصوصية	هل الجمع ضروري، والوصول محدود، والإفصاح متوقع، والاحتفاظ مضبوط؟
المساءلة	هل يمكن تفسير القرارات، والإصدارات، والاستثناءات، والحوادث، والتغييرات، ومراجعتها بصورة مستقلة؟

إطار لاتخاذ قرار أخلاقي

1. عرّف القرار والأشخاص المتأثرين والغرض المقصود.
2. حدد المعايير المهنية والقوانين والعقود والسياسات وتوقعات المشاركين المعقولة ذات الصلة.
3. اسرد المنافع والأضرار وعدم اليقين وتعارض المصالح لكل مجموعة من أصحاب العلاقة.
4. اختر الضرورة، والتناسب، والصلة بالوظيفة، وإمكانية الوصول، والخصوصية، والبدائل الأقل ضرراً.
5. اطلب مراجعة مستقلة من التخصصات والمجموعات المناسبة لمستوى المخاطر.
6. اختر ووثق الخيار الذي يمكن الدفاع عنه علناً وأمام الشخص المتأثر.
7. عرّف التواصل، والموافقة أو الإشعار، والمراجعة البشرية، والاعتراض، والمعالجة، والمراقبة، والخروج.
8. أعد التقييم بعد ظهور أدلة جديدة، أو حوادث، أو تغيير تنظيمي، أو تطور قانوني.

موضوعات قانونية عبر الولايات القضائية

تختلف قوانين التوظيف، والمساواة، والإعاقة، وحماية البيانات، وحماية المستهلك، والعمل، والممارسة المهنية، والذكاء الاصطناعي. والنقاط التالية ليست رأياً قانونياً؛ بل توضح لماذا يتطلب التقييم عالي المخاطر أو المبتكر مراجعة قانونية ومتخصصة محلية.

الموضوع	الآثار على التقييم
الصلة بالوظيفة والضرورة	ينبغي أن ترتبط متطلبات الاختيار بالعمل وأن تدعمها أدلة تحقق مناسبة. يصعب الدفاع عن العوائق غير الضرورية.
التمييز والنتائج السلبية	قد تنظم القوانين المعاملة المباشرة، والآثار غير المباشرة، والتحرش، والانتقام، والخصائص المحمية. افحص التصميم والنتائج معاً.
الإعاقة والتيسيرات	قد تكون التعديلات المعقولة أو البدائل مطلوبة. تجنب الاستفسارات الطبية أو استنتاجات الصحة خارج المراحل والأغراض القانونية.
حماية البيانات والخصوصية	عرّف الغرض والأساس القانوني، وقلل البيانات، وقدم المعلومات، واضبط الوصول والاحتفاظ، وأدر الموردّين والتحويلات.
القرارات المؤتمتة والذكاء الاصطناعي	قد تنطبق متطلبات إشعار، أو تقييم أثر، أو مراجعة بشرية، أو تدقيق تحيز، أو توثيق، أو تفسير، أو قيود حسب الولاية والنظام.
حفظ السجلات والتدقيق	قد يحتاج أصحاب العمل إلى أدلة على اتساق الاختيار والصلة بالوظيفة مع الالتزام أيضاً بقواعد تقليل البيانات والاحتفاظ.
الألقاب المهنية والأدوات المقيدة	تنظم بعض الولايات لقب «أخصائي نفسي»، أو النشاط السريري، أو الوصول إلى أدوات محددة.
الحقوق الجماعية وحقوق الموظفين	قد تنطبق المشاورات، ومجالس العمل، والنقابات، ومراقبة مكان العمل، وحقوق ممثلي الموظفين.

سياق توضيحي: الولايات المتحدة

في الولايات المتحدة قد تكون أحكام Title VII، وقانون الأمريكيين ذوي الإعاقة، وقانون التمييز في التوظيف بسبب السن، والإرشادات الموحدة لإجراءات اختيار الموظفين، ومتطلبات الولايات أو السلطات المحلية ذات صلة. توضح إرشادات EEOC الرسمية أن الاختبارات وإجراءات الاختيار قد تكون مفيدة، لكنها قد تصبح غير قانونية عندما تميز، أو تستبعد بشكل غير متناسب مجموعات محمية من دون مبرر كاف مرتبط بالوظيفة، أو تستبعد أشخاصاً ذوي إعاقة من دون استيفاء المتطلبات ذات الصلة والنظر في التيسيرات. وقد تضيف القواعد المحلية بشأن أدوات التوظيف المؤتمتة التزامات تتعلق بالإشعار، أو التدقيق، أو الموافقة، أو العملية البديلة، أو الاحتفاظ.

قاعدة الأربعة أحماس أداة فحص عملية ضمن سياق الإرشادات الموحدة، وليست «منطقة أمان» قانونية عالمية. وتظل المؤسسة بحاجة إلى فحص حجم العينة، والحجم العملي للفروق، والأدلة الإحصائية، والصلة بالوظيفة، والإجراءات البديلة، والقانون المنطبق.

سياق توضيحي: المملكة المتحدة

في المملكة المتحدة، تُعد أحكام Equality Act 2010، وقانون حماية البيانات البريطاني، وقانون العمل، والالتزامات المهنية ذات صلة. ينبغي لأصحاب العمل مراعاة التمييز المباشر وغير المباشر، والتعديلات المعقولة، والتصميم القابل للوصول، والمعالجة القانونية والشفافة، وتقليل البيانات، والقرارات المؤتمتة، والاحتفاظ، وحقوق الأفراد. وقد نشر مكتب مفوض المعلومات ICO إرشادات حول التوظيف والذكاء الاصطناعي، وأشار إلى نقاط ضعف في الخصوصية، والدقة، والعدالة، والشفافية، والحوكمة في أدوات التوظيف.

سياق توضيحي: الاتحاد الأوروبي

قد تنطبق اللائحة العامة لحماية البيانات GDPR، وقوانين التوظيف والمساواة الوطنية، وقانون الذكاء الاصطناعي الأوروبي. يصف قانون الذكاء الاصطناعي بعض الأنظمة المستخدمة في التوظيف والاختيار وقرارات العمل والترقية وتخصيص المهام أو مراقبة الأداء على أنها عالية المخاطر، وفق تعريفاته واستثناءاته ومراحل تطبيقه.

هذا الإصدار مؤرخ في 17 يوليو 2026. ولا تزال جداول تطبيق تشريعات الاتحاد الأوروبي والإرشادات التنفيذية مجالات نشطة للتغيير، بما في ذلك أعمال تشريعية مرتبطة بالتبسيط وتأخير تطبيق بعض قواعد الأنظمة عالية المخاطر. يجب على أي مؤسسة تستخدم الذكاء الاصطناعي في التوظيف أن تتحقق من النص القانوني النهائي وأحدث إرشادات المفوضية الأوروبية والسلطات الوطنية، لا أن تعتمد على ملخص ثابت.

حوكمة insyt2: افصل بين الأسئلة العلمية والقانونية والأخلاقية

قد يكون الأسلوب قابلاً للدفاع علمياً ومع ذلك غير قانوني أو غير أخلاقي في تنفيذ محدد؛ وقد تكون العملية قانونية لكنها ضعيفة علمياً. كل سؤال يحتاج إلى الخبرة المناسبة.

المشاركة المستنيرة في سياقات العمل

قد تكون قدرة المرشحين والموظفين على رفض التقييم محدودة، لذلك لا يلغي توقيع نموذج موافقة اختلال ميزان القوة. قدم معلومات واضحة ومحددة أياً كان الأساس القانوني. لا تدمج البحث الاختياري، أو تطوير المنتج، أو التسويق في مشاركة إلزامية في التوظيف. اشرح بدقة عواقب عدم المشاركة، والبدائل، والمستلمين، والاحتفاظ، والتسجيل، والأتمتة، ومسارات المراجعة.

النطاق المهني والألقاب المحمية

قد تدعم دورة قصيرة من ناشر ما الاستخدام الكفاء لأداة محددة، لكنها لا تمنح لقباً محمياً مثل «أخصائي نفسي»، ولا صلاحية تشخيص الحالات الصحية، ولا كفاءة عامة في تطوير الاختبارات. تختلف المتطلبات حسب الولاية والأداة. يجب أن يستطيع الممارس تحديد ما هو مؤهل للقيام به تحديداً، وكيف تم تقييم كفاءته، وأين ينطبق الإشراف، ومتى يلزم الإحالة.

تعارض المصالح

- افصح عندما يكون المستشار بائعاً للأداة الموصى بها أو يحصل على عمولة منها.
- افصل المراجعة الفنية عن الأهداف التجارية كلما أمكن.
- لا تسمح للمؤزد بأن يقيم أدلته عالية المخاطر بنفسه من دون تحيد مستقل.
- احيم الممارسين الذين يثيرون مخاوف حول الصدق، أو العدالة، أو الخصوصية، أو إمكانية الوصول، أو الأمن.
- وثق الضغوط لتغيير العتبات، أو إخفاء النتائج، أو نشر الأداة خارج الاستخدام المدعوم.
- لا تستخدم معلومات المشاركين السرية لأهداف تنظيمية أو تجارية غير مرتبطة بالغرض.

سجل أخلاقي عملي

في المشروعات عالية المخاطر، احتفظ بسجل أخلاقي مختصر إلى جانب وثائق التحقق والقانون. سجل المنفعة المقصودة، والأشخاص المتأثرين، والأضرار المحتملة، والبدائل التي نوقشت، والمشاورات، والأدلة، والقرارات، ووجهات النظر المخالفة، والمخاطر المتبقية، والتواصل، والمعالجة، وتاريخ المراجعة. يخلق ذلك منطقاً قابلاً للتتبع بدلاً من عبارة عامة بأن «الأخلاقيات أخذت في الاعتبار».

متى تتوقف أو تؤجل أو تحيل؟

الموقف	الاستجابة المسؤولة
التفسير المطلوب يدخل في المجال السريري أو الطبي أو التشخيصي.	أوقف التفسير وأجل إلى مهني مؤهل ومصرح له بالشكل المناسب.
صاحب علاقة يطلب استنتاجاً محمياً أو حساساً لا يرتبط بالعمل الجوهري.	تحذّر الضرورة والقانونية، ووثق الطلب، وصعد عبر الحوكمة الأخلاقية أو القانونية.
الأدلة الفنية لا تدعم الاستخدام عالي المخاطر المقترح.	حد من الاستخدام، أو أعد التصميم، أو جرب في وضع ظل، أو لا تنشره.
المشارك يتعرض لضيق أو يكشف مخاوف تتعلق بالحماية.	استجب بإنسانية، وحافظ على حدود الدور، واتبع بروتوكول الحماية والإحالة في المؤسسة.
ضغط تجاري يتعارض مع الأدلة أو حماية المشاركين.	سجل التعارض، واطلب مراجعة مستقلة، ورفض تقديم ادعاء أو توصية مضللة.
تفتقر إلى الكفاءة للتحليل أو القرار.	اطلب إشرافاً أو أجل العمل بدلاً من التعلم عبر حالة عالية المخاطر بلا دعم.

السجلات المهنية

احتفظ بسجلات متناسبة مع مستوى المخاطر ومتطلبات الاحتفاظ القانونية: الغرض والنطاق، والأداة والإصدار، والإدارة والتيسيرات، والدرجة وفحوص الجودة، ومنطق التفسير أو القرار، وكفاءة المستخدم، وتواصل المشارك، والحوادث، والمراجعات، والاستثناءات، والموافقات. افصل الملاحظات المهنية عن التعليقات التخمينية. اكتب كل سجل على افتراض أن الشخص المتأثر، أو المدقق، أو الجهة التنظيمية، أو ممارساً كفؤاً آخر قد يحتاج إلى فهمه.

الفصل

19

الذكاء الاصطناعي، والخوارزميات، والألعاب،
والفيديو، والتقييم الرقمي

تخلق الواجهات الجديدة مخاطر جديدة، لكن المتطلبات الأساسية تظل الصلة بالوظيفة، والصدق، والعدالة، والشفافية، والمساءلة.

ما الذي يتغير وما الذي لا يتغير؟

يمكن للذكاء الاصطناعي دعم إنشاء المفردات، والتقديم التكيفي، وتسجيل اللغة الطبيعية، وبيانات الألعاب، ونماذج التوصية، وكتابة التقارير، والمراقبة، وأتمتة سير العمل. قد تزيد هذه القدرات من التوسع والاستجابة، لكنها قد تحجب البنية، وتجمع بيانات مفرطة، وتعيد إنتاج قرارات تاريخية، وتنحرف بمرور الوقت، وتشجع تحيز الأتمتة.

الواجهة الجديدة لا تخفض معيار الدليل. تؤكد الإرشادات المهنية بشأن الاختيار المدعوم بالذكاء الاصطناعي أن الأساليب الجديدة يجب أن تخضع للمراجعة نفسها التي تخضع لها الإجراءات التقليدية، مع ضوابط إضافية لمصدر البيانات، ومعنى الخصائص، وتطوير النموذج، وقابلية التفسير، والانحراف، والأمن، والإشراف البشري.

سلسلة التقييم الخوارزمية

الطبقة	الأسئلة الأساسية
مصدر البيانات	من أنتج البيانات؟ ولأي غرض أصلي؟ وتحت أي فرص وإشعار؟ وما البيانات المفقودة أو التحيز التاريخي الموجود؟
الخصائص	ما السلوك أو المتغير الذي يدخل في الدرجة؟ هل هو ذو صلة بالعمل، وضروري، ومستقر، وقابل للفهم؟ وهل يمكن أن يعمل كمؤشر بديل لخصائص حساسة؟
التسميات والمعايير	ما النتيجة التي يجري التنبؤ بها؟ هل هي ثابتة، وغير ملوثة، ومناسبة زمنياً، ومرغوبة أخلاقياً؟ وهل تحمل تحيزات المديرين أو الأنظمة التاريخية؟
تطوير النموذج	كيف فصلت بيانات التدريب والضبط والاختبار؟ وكيف عولج التسرب، وفرط الملاءمة، وعدم التوازن، وعدم اليقين؟
الأداء	إلى أي مدى يعمم النموذج، ويتعاير، ويضيف قيمة؟ وكيف يختلف الأداء حسب المجموعة والسياق ونطاق الدرجة؟
دمج القرار	هل المخرج توصية، أو ترتيباً، أو عتبة، أو إجراء مؤتمتاً؟ وهل يستطيع الأشخاص فهمه، والسؤال عنه، وتصحيحه؟
النشر	كيف تُحكم الإصدارات، والانحراف، وتحولات البيانات، والهجمات، وإمكانية الوصول، والحوادث، والشكاوى، وإعادة التدريب، والإيقاف؟

التقييم القائم على الألعاب

يمكن للمهام الشبيهة بالألعاب أن تولد ملاحظات سلوكية متكررة وواجهة جذابة، لكن الجاذبية ليست صدقاً. قد يعكس سلوك اللعبة نفسه الاستراتيجية، أو السرعة الحركية، أو الإلمام بالألعاب، أو أداء الجهاز، أو تفضيل المخاطر، أو القراءة، أو الانتباه، أو البنية المقصودة. يجب على الموزع تعريف سلسلة الدليل من السلوك إلى الخاصية إلى الدرجة إلى الاستنتاج الوظيفي، ودراسة آثار الجهاز وإمكانية الوصول، وإزالة المتطلبات الحسية أو الحركية غير ذات الصلة، وشرح ما إذا كانت الصعوبة التكيفية تغير معنى الدرجة.

اسأل ما إذا كانت اللعبة مجرد صيغة تقديم لبنية معروفة، أو مهمة جديدة ذات نموذج أدلة محدد، أو مولداً غامضاً للخصائص. الفئة الثالثة تحمل أكبر مخاطر الاختصارات المستندة إلى البيانات وضعف قابلية التفسير.

تسجيل اللغة الطبيعية

قد يستخدم التسجيل الآلي للاستجابات المكتوبة أو المنطوقة المحتوى، والبنية، والمفردات، والتشابه، والمشاعر، والخصائص الصوتية، أو أنماطاً متعلمة. النموذج المدرب على تقليد المقيمين البشر يرث جودة تلك التقديرات وتحيزاتها، وقد يكتشف اختصارات غير مرتبطة بالمتطلب المستهدف. تحقق منه مقابل بُنى محددة بوضوح ومعايير ذات صلة، وادرس الاستجابات غير المعيارية والهجومية، وافحص الأداء حسب اللغة واللهجة، ووفر مراجعة بشرية.

يغير الذكاء الاصطناعي التوليدي أيضاً بيئة الاستجابة. قرر ما إذا كان استخدام الأدوات مسموحاً أو محظوراً أو جزءاً من المهمة. قد تسمح عينة عمل واقعية بمساعد ذكي بينما تقيس تعريف المشكلة، والتحقق، والحكم، والمساءلة. أما مهمة أساسية في الكتابة أو الاستدلال فقد تحتاج إلى شروط مضبوطة. اذكر القاعدة، وصمّم حولها، واستخدم تحققاً لاحقاً عند الحاجة.

الفيديو والصوت والقياسات الحيوية واستنتاج الانفعال

يمكن للفيديو والصوت أن يكونا وسائل مفيدة لتقديم أو تسجيل استجابات مرتبطة بالعمل. أما استخراج حركة الوجه، أو اتجاه النظر، أو الخصائص الصوتية، أو الانفعال المستنتج فيخلق مخاطر كبيرة تتعلق بالصدق، والإعاقة، والثقافة، والخصوصية، والقانون. يتأثر السلوك الظاهر بالتنوع العصبي، والإعاقة، واللغة، والمعدات، والقلق، والإضاءة، والثقافة. لا تفترض أن التعبير الظاهر يكشف الصدق، أو الارتباط، أو الشخصية، أو الانفعال. بعض الولايات تحظر أو تقيد استنتاج الانفعالات أو استخدام القياسات الحيوية في العمل. استخدم أقل الأساليب تدخلاً واطلب مراجعة متخصصة.

إطار تدقيق الذكاء الاصطناعي

لن تقوم insyt2 أبداً بـ...

اعتبار الحداثة أو الأتمتة أو تعقيد النموذج دليلاً على الجودة.

في insyt2 نؤمن بأن...

مجال التدقيق	الحد الأدنى من الأسئلة
الاستخدام المقصود	ما القرار والفئة المدعومان؟ ما الاستخدامات المحظورة؟ وهل مكون الذكاء الاصطناعي ضروري؟
البنية والخصائص	ما الذي تمثله كل خاصية جوهرية؟ وما أدلة تحليل الوظيفة التي تدعماها؟
مصدر البيانات	هل بيانات التدريب والتقييم قانونية وممثلة وحديثة ومنفصلة عن تسرب النتائج؟
الأداء الفني	ما الخطأ، والمعايرة، وعدم اليقين، والمتانة، والأداء حسب المجموعة والسياق؟
العدالة والوصول	ما تحليلات القياس، والتنمؤ، والنتائج، وإمكانية الوصول، والمؤشرات البديلة التي أنجزت؟ وما البدائل التي اختُبرت؟
الشفافية	هل يمكن للعميل، والممارس، والمراجع، والشخص المتأثر الحصول على تفسير يناسب دوره؟
الإشراف البشري	هل يستطيع المراجعون اكتشاف الخطأ، والوصول إلى الأدلة، وتصحيح البيانات، والتجاوز وفق قواعد، وتجنب الموافقة الشكلية؟
الأمن وسوء الاستخدام	هل يمكن التحايل على النظام، أو استخراج النموذج، أو تسميمه، أو التلاعب به؟ وهل المحتوى والمدخلات والبيانات والنماذج محمية؟
التغيير والانحراف	ما الذي يطلق إعادة التحقق، أو إعادة التدريب، أو الإشعار، أو التراجع، أو التعليق، أو الإيقاف؟
الحوكمة	من المسؤول عبر المورد والجهة المستخدمة؟ وهل حقوق التدقيق والسجلات والوثائق والتزامات الحوادث قابلة للإنفاذ؟

العدالة متعددة المقاييس وخاصة بالقرار

لا توجد إحصائية واحدة للعدالة تحسم كل اختيار قيمي. قد تتعارض مساواة معدلات الاختيار، ومعدلات الخطأ، والمعايرة، وتكافؤ الفرص، والعدالة الفردية، خصوصاً عندما تختلف المعدلات الأساسية أو يكون المعيار غير كامل. اختر المقاييس بالارتباط بالقرار والضرر والسياق القانوني والنموذج السيكمترية. اذكر المفاضلات واختبر ما إذا كانت البدائل تستطيع الحفاظ على قيمة القرار مع تقليل الفروق الضارة.

لا تعدل الدرجة آلياً لتحقيق مقياس تكافؤ مفضل من دون مراعاة صدق البنية، والمعايرة، والإيجابيات الكاذبة، والسلبيات الكاذبة، والآثار اللاحقة. تدخل العدالة نفسه قاعدة قرار تحتاج إلى تحقق وحوكمة.

إشراف بشري ذو معنى

- يفهم المراجعون ما الذي يقيسه النموذج، وكيف تم تقييمه، وأين يفشل.
- تعرض الواجهة عدم اليقين، والأدلة المصدرة، والإشارات، والقيود بدلاً من رقم واحد يبدو نهائياً.
- لدى المراجعين وقت وصلاحيّة وأدلة بديلة كافية لمساءلة المخرج.
- أسباب التصحيح والتجاوز المسموح بها محددة ومسجلة.
- تُراقب أنماط القبول والتجاوز والخطأ حسب المراجع، والدور، والموقع، والمجموعة ذات الصلة.
- يستطيع الأشخاص المتأثرون تصحيح البيانات الجوهرية والحصول على مراجعة ذات معنى.
- يمكن إيقاف النظام، أو التراجع إلى إصدار سابق، أو استبداله عندما تفشل الأدلة أو الضوابط.

إشارة تحذير لدى insyt2: التنبؤ الغامض قد يخفي الإشارة الخطأ

قد تتنبأ خاصية مملوكة لأنها تلتقط امتيازاً تاريخياً، أو مؤشراً بديلاً لفئة محمية، أو تسرباً من المعيار، أو سلوك المنصة. القوة التنبؤية وحدها ليست كافية.

استخدام الذكاء الاصطناعي التوليدي في أعمال التقييم

يمكن للذكاء الاصطناعي التوليدي المساعدة في صياغة أسئلة تحليل الوظيفة، وبدائل المفردات، ومؤشرات التسجيل، والتقارير، والشفرة، والوثائق، وتواصل المشاركين. وينبغي أن يعمل داخل سير عمل مضبوط:

1. لا تضع مفردات مقيدة، أو بيانات سرية، أو بيانات شخصية في أدوات غير معتمدة.
2. ابدأ بمواصفات خبيرة وبُنى مقصودة واضحة.
3. تحقق من الحقائق، والمصادر، وحقوق النشر، واللغة، والمحتوى الثقافي، وإمكانية الوصول.
4. أخضع المفردات أو القواعد المولدة لمراجعة الخبراء والتجريب نفسه الذي تخضع له المواد البشرية.
5. سجل النموذج، والإصدار، والتعليمات، والمراجعين، والتغييرات، والموافقة النهائية عندما تكون جوهرية.
6. اختبر الهلوسة، والتحيز، والتسرب، والأمن، والسلوك غير المستقر.
7. أبقِ إنساناً مؤهلاً مسؤولاً عن كل تفسير للدرجات وكل ادعاء متعلق بالقرار.

أنماط فشل خاصة بالذكاء الاصطناعي وضوابطها

نمط الفشل	كيف قد يظهر؟	الحد الأدنى للضبط
تسرب المعيار	يبدو النموذج دقيقاً بصورة غير معتادة لأن معلومات ما بعد القرار أو معلومات النتيجة دخلت في الخصائص.	مراجعة بيانات تراعي الزمن، وتتبع أصل الخصائص، وعينة اختبار مستقلة، وإعادة إنتاج عند نقطة القرار الحقيقية.
التمييز عبر مؤشر بديل	يعمل الموقع، أو الجهاز، أو نمط اللغة، أو إشارة التعليم، أو خاصية سلوكية كمؤشر بديل لحالة محمية أو حساسة.	مراجعة البنى، وتحليل المؤشرات البديلة، والأداء حسب المجموعات، واختبار خصائص بديلة، والإزالة عندما لا تكون ضرورية.
تحيز التسميات التاريخية	يتعلم النموذج تقييمات مديرين غير متنسقة، أو تفضيلات توظيف سابقة، أو تفاوتاً في الفرص.	تدقيق المعيار، ودراسة جودة المقيمين، ونتائج بديلة، وإعادة وزن أو إعادة تصميم، ومساءلة بشرية.
فرط الملاءمة	دقة عالية أثناء التطوير تنخفض بشدة في عينة أو وظيفة أو موقع أو فترة جديدة.	فصل صارم بين التدريب والضبط والاختبار، وتحقق متقاطع، وتحقق خارجي، وتفضيل البساطة، ومراقبة بعد النشر.
فشل المعايرة	الاحتمالية أو نطاق الدرجة لا يطابق النتائج المرصودة، خصوصاً في بعض المجموعات.	مخططات معايرة، وتحليل مجموعات فرعية، وإعادة معايرة تحت حوكمة، ولغة حذرة عند التعبير عن الاحتمالات.
تحيز الأتمتة	يتبع المراجعون النموذج حتى عند وجود دليل وظيفي مناقض.	تصميم واجهة يعرض عدم اليقين، وتدريب، وحكم مستقل قبل إظهار النموذج، ومراقبة الاستثناءات.
فشل إمكانية الوصول	تضر الواجهة أو الخاصية بالأشخاص الذين يستخدمون تقنيات مساعدة أو ذوي اختلافات حسية أو حركية أو لغوية أو معرفية.	اختبار إمكانية الوصول مع المستخدمين، وتيسيرات، وأساليب بديلة، ومراجعة البنية على مستوى الخاصية.
تغيير صامت للنموذج	يحدث النموذج أو خط معالجة الخصائص أو التعليمات أو المعايير أو العتبة من دون إعادة تحقق لدى العميل.	تنبيه الإصدار، وإشعار تغيير، وسجلات تدقيق، وبوابات موافقة، واختبار رجعي، وإمكانية التراجع.
التحاييل أو الاستجابة الهجومية	يتعلم المرشحون اختصارات، أو يولدون استجابات، أو يتلاعبون بالبيانات السلوكية بطرق لا ترتبط بالبنية.	نمذجة التهديدات، ومهام أصيلة، والتحقق، ومراقبة الانكشاف، وإعادة التصميم بدلاً من اللجوء افتراضياً إلى مراقبة متطفلة.
حلقة تغذية راجعة	القرارات التي يصنعها النموذج تشكل البيانات المستقبلية وتجعل النموذج يبدو كأنه يؤكد نفسه.	راقب من يحصل على الفرص، واستخدم أدلة مضادة للواقع أو تجريبية حيث أمكن، وراجع مناطق العمى لدى المرشحين المرفوضين.

قرار: اعتماد، اعتماد مشروط، أو عدم اعتماد

ينبغي لقرار الموافقة المسؤول أن يحدد إصدار النظام بالضبط، والغرض، والفئة، والبيانات، ودور المخرج في القرار، والشروط، والمراقبة، وتاريخ انتهاء الموافقة. عبارة «ذكاء اصطناعي معتمد» أوسع من اللازم. قد يقيد الاعتماد المشروط النموذج بوضع ظل، أو بتوصية لا تستطيع رفض المرشحين، أو بعائلة وظيفية محددة، أو بفترة تقييم قصيرة. ويكون قرار عدم الاعتماد مناسباً عندما تكون البنية غير واضحة، أو لا يمكن إثبات مصدر البيانات، أو الصلة بالوظيفة ضعيفة، أو أدلة المجموعات الفرعية أو إمكانية الوصول غير كافية، أو المراجعة البشرية الحقيقية غير ممكنة، أو يرفض المورد دعم التدقيق وضبط التغيير.

قاعدة insyt2 للموافقة: اعتمد إصداراً واستخداماً، لا علامة تجارية

يجب أن تحدد الموافقة الفئة، والغرض، والإصدار، والتسجيل، والضوابط، والأدلة، وتاريخ المراجعة، وأن تنتهي عندما تتغير هذه الشروط.

الفصل

20

حوكمة التقييم، ومركز التميز، والنضج

تحتاج المؤسسات إلى نظام تشغيل لجودة التقييم، لا إلى خبراء متفرقين وعقود موردين منفصلة.

لماذا نحتاج إلى مركز تميز؟

في insyt2 نؤمن بأن...

لن نقوم insyt2 أبداً بـ...

السماح لفرق منفصلة بشراء تقييمات أو بنائها أو نشرها من دون معايير مشتركة أو ملكية واضحة لدورة الحياة.

غالباً ما يتوزع التقييم بين التوظيف، وتطوير القيادات، والتعاقب، وبرامج الخريجين، وتحليلات الأفراد، وتقنيات الموارد البشرية، والوحدات المحلية، والموردين الخارجيين. قد تقيس فرق مختلفة بُنى متداخلة، أو تخزن بيانات حساسة، أو تستخدم معايير متعارضة، أو تعيد استخدام أدوات تطويرية للاختيار. يمكن لمركز تميز للتقييم أن يضع المعايير، ويحافظ على المخزون، ويعتمد الاستخدامات عالية المخاطر، ويدرب المستخدمين، ويقدم الاستشارة، ويراقب النتائج، ويبنى أساليب قابلة لإعادة الاستخدام.

يمكن أن يكون النموذج مركزياً أو اتحادياً. في النموذج المركزي، يملك فريق متخصص معظم التصميم والاعتماد. وفي النموذج الاتحادي، يعمل ممارسون محليون مؤهلون ضمن سياسة مشتركة ويتلقون ضماناً مركزياً. وفي الحالتين يجب أن تكون حقوق القرار ومسارات التصعيد صريحة.

المكونات الأساسية للحوكمة

المكون	الحد الأدنى للمحتوى
السياسة	الغرض، والمبادئ، والنطاق، وفئات المخاطر، والكفاءة، والاستخدامات المعتمدة، وحقوق المشاركين، وقواعد البيانات والذكاء الاصطناعي، والاستثناءات، والتدقيق، والعقوبات.
مخزون التقييمات	الأداة، والمالك، والغرض، والفئة، ودور القرار، والمورد، والإصدار، والمعايير، ومكان البيانات، والمخاطر، والأدلة، وتاريخ المراجعة.
سير عمل الاعتماد	فحص المخاطر، ومتطلبات الأدلة، والمراجعات المستقلة، والتجريب، وبوابة الإطلاق، واعتماد التغيير، والتعليق، والخروج.
المعايير والقوالب	تحليل الوظيفة، والمخطط، وطلب العروض RFP، والتحقق، والإشعار، والتيسيرات، وضبط جودة التسجيل، والمراقبة، والحوادث، والتدقيق.
نظام الكفاءة	ملفات الأدوار، والمؤهلات، والتدريب، والممارسة الخاضعة للإشراف، والمعايرة، والتطوير المستمر، والتحكم في الوصول.
المراقبة والضمان	لوحات متابعة، وتحليل مجموعات فرعية، وانحراف، وتجربة، وتدقيق، ومراجعة المورد، ومراقبة النموذج والإصدار.
المعرفة والدعم	إرشادات، وساعات استشارة، ومكتبة حالات، ومجتمعات ممارسة، وتحديثات بحثية.
منتدى الحوكمة	مدخلات من الأعمال، والتقييم، والقانون، والخصوصية، والأمن، وإمكانية الوصول، والتنوع والشمول، والتكنولوجيا، والمشتريات، والموظفين، والمخاطر بحسب الحاجة.

حقوق مخزون التقييمات

- الأداة، والدرجة، والمورد، ومالك العقد، والمالك العلمي، وجهة الاتصال التقنية.
- الغرض، والفئة المستهدفة، والدول، واللغات، والحجم، ومستوى المخاطر، ودور التقييم في القرار.
- البنى، والأساليب، والوقت، ونمط التقديم، والتيسيرات، وتواصل المشاركين.
- إصدار المفردة أو النموذج، ونموذج التسجيل، والمعيان، والعتبة، والمركب، والتقرير.
- ملخص الأدلة: تحليل الوظيفة، والثبات، والصدق، والعدالة، وإمكانية الوصول، والدراسات المحلية.
- فئات البيانات، والأنظمة، والمعالجون الفرعيون، والدول، والاحتفاظ، والوصول، والحذف.
- متطلبات التدريب والتأهيل، والمستخدمون المعتمدون، وتواريخ الانتهاء.
- مؤشرات المراقبة، وآخر مراجعة، والحوادث، والقيود، والشروط، والمراجعة التالية.
- الحالة: معتمد، أو مشروط، أو معلق، أو قديم، أو متقاعد، مع ذكر الأسباب.

نموذج نضج من خمسة مستويات

المستوى	الخصائص	الأولوية للانتقال
1. ارتجالي	مشتريات محلية، واستخدامات غير معروفة، ولا مخزون، وقبول التقارير كما هي.	حصر الأدوات ووقف الاستخدامات عالية المخاطر غير المدعومة.
2. مضبوط	سياسة أساسية، ومراجعة مشتريات، ومشرفون مدربون، ومخزون جزئي.	تعريف فئات المخاطر، وبوابات الأدلة، والملاك، والتيسيرات، وخرائط البيانات.
3. موحد	أطر مشتركة، وموردون معتمدون، وقوالب، وتدريب، وبوابات إطلاق، ومراقبة.	تحسين التحقق المحلي، وتحليل المجموعات الفرعية، والتدقيق، وضبط التغيير.
4. قائم على الأدلة	بيانات متكاملة، ودراسات نتائج، وضمان مستقل، ومركز تميز ناضج، وكفاء اصطناعي محكوم.	تحسين المنفعة، والعدالة، والتجربة، والأدلة عبر الأسواق.
5. تكيفي وشفاف	مراقبة مستمرة، وتعلم سريع مضبوط، وشفافية محورها المشارك، وثقافة تحدٍ قوية.	استدامة الكفاءة، واستشراف الأساليب الناشئة، ومشاركة الممارسات المتحقق منها.

مبادئ سياسة تستحق النص عليها

- لا تقييم من دون غرض محدد ومشروع ومتناسب.
- لا استخدام عالي المخاطر من دون أدلة على الصلة بالوظيفة، والصدق، والثبات، والعدالة، وإمكانية الوصول، وقواعد القرار.
- لا تغيير جوهري من دون مراجعة أثر وضبط للإصدار.
- لا إعادة استخدام خفية لبيانات التطوير أو البحث أو مراقبة المساواة أو المرشحين.
- لا قرار سلبي مؤتمت من دون مراجعة بشرية قانونية وذات معنى وكفاءة عندما تكون مطلوبة.
- لا وصول للممارس يتجاوز الكفاءة المثبتة والحاجة المهنية.
- لا قبول لادعاء مورد من دون أدلة قابلة للفحص ومساءلة تعاقدية.
- لا لوحة معلومات من دون ملاك وعتبات وإجراءات تصحيحية.
- لا نظام من دون خطة تقاعد وخروج للبيانات.

دورة التدقيق

1. أكد المخزون والنطاق والعينة القائمة على المخاطر.
2. راجع الغرض، وتحليل العمل، والأدلة، وشروط الاستخدام، وقواعد القرار، والكفاءة، والتواصل، وإمكانية الوصول، والبيانات، وضوابط المورد.
3. أعد إنتاج درجات مختارة وتتبع القرارات من الأدلة الخام إلى النتيجة.

4. حل المراقبة، ونتائج المجموعات الفرعية، والاستثناءات، والحوادث، والتغييرات، والشكاوى، والإجراءات التصحيحية.

5. صُفِّ النتائج حسب الأثر والاستعجال، وعيِّن المسؤولين والمواعيد. 6. تحقق من إغلاق الإجراءات، وحَدِّث السياسة أو التدريب أو العقود أو المخزون أو تصميم النظام. 7. ارفع الموضوعات والمخاطر المتبقية إلى منتدى الحوكمة المسؤول. منظور insyt2 للمسيرة المهنية: الحوكمة تخصص مهني قائم بذاته تناسب حوكمة التقييم متخصصي الموارد البشرية ذوي الخبرة الذين يجمعون بين الثقافة العلمية، والحكم على المخاطر، والتأثير في أصحاب العلاقة، وتصميم العمليات.

بناء مسيرة مهنية في الاختبارات السيكومترية

مسار عملي من متخصص موارد بشرية إلى أخصائي تقييم موثوق.



في هذا الجزء

21 المسارات المهنية واختبارات الأدوار

22 نموذج كفاءة ممارس التقييم

23 المؤهلات والتعلم المتعمد

24 خارطة انتقال لمدة اثني عشر شهراً

25 المحفظة المهنية ومسارات التوظيف والاستشارات

العلم • العدالة • جودة القرار • الحكم المهني

رؤية insyt2

السؤال ليس أي اختبار يبدو أكثر إبهاراً، بل أي دليل يخدم القرار، والأشخاص المتأثرين به، وما يحدث بعد ذلك بصورة أفضل.

21

الفصل

المسارات المهنية واختيارات الأدوار

لا توجد مسيرة مهنية واحدة في السيكمتريا؛ اختر العمل الذي تريد أن تصبح موثقاً في أدائه.

ابدأ بالعمل، لا بالمسمى

في insyt2 نؤمن بأن...

لن تقوم insyt2 أبداً بـ...

تشجيع شخص على ادعاء خبرة سيكمترية قبل أن يستطيع إثبات قدرات القياس والتقييم والحوكمة الأساسية.

المسار	العمل المعتاد	مدخل شائع	التطور الأطول مدى
عمليات التقييم أو التنسيق	رحلة المرشح، وإدارة المنصة، والدعم، والسجلات، والحوادث، واللوجستيات.	عمليات التوظيف، أو خدمات الموارد البشرية، أو تنسيق مراكز التقييم.	قائد عمليات، أو استشاري تنفيذ، أو عمليات المنتج.
مستخدم اختبارات أو ممارس مؤهل	اختيار الأداة، والإدارة، والتفسير، والتغذية الراجعة، وإرشاد المديرين.	الموارد البشرية، أو التعلم والتطوير، أو التدريب، أو استقطاب المواهب مع تدريب معترف به لاستخدام الاختبارات.	ممارس أول، أو مقيم قيادات، أو استشاري مواهب.
أخصائي تقييم المواهب	تحليل الوظيفة، وتصميم الاختيار، والمقابلات المنظمة، والمحاكاة، والتحقق، والمراقبة.	استقطاب المواهب، أو تحليلات الأفراد، أو علم النفس المهني، أو مزود تقييمات.	قائد تقييم، أو رئيس مركز تميز، أو مدير استشارات.
مصمم تقييم أو استشاري	نماذج الكفاءات، واختبارات الحكم الموقفي، والتمارين، والاستبيانات، وأدلة التسجيل، وحلول العملاء.	الاستشارات، أو تصميم التعلم، أو علم النفس المهني، أو البحث التطبيقي.	استشاري رئيسي، أو قائد ممارسة، أو مستشار مستقل.
سيكمترى أو عالم قياس	تطوير المقاييس، والنماذج العالمية، وIRT، والربط، وDIF، والتحقق، والبحث الإحصائي.	دراسات عليا كمية في علم النفس، أو السيكمتريا، أو الإحصاء، أو القياس التربوي.	سيكمترى أول، أو مدير بحوث، أو كبير العلماء.
محلل أفراد أو عالم اختيار	دراسات تنبؤية، ومنفعة، وتحليل عدالة، وتجارب، ولوحات متابعة، ونماذج قوى عاملة.	تحليلات الأفراد، أو علم البيانات، أو بحوث I-O.	رئيس تحليلات التقييم أو قائد علوم القوى العاملة.
مدير منتج تقييم	استراتيجية المنتج، والمتطلبات، وتجربة المستخدم، والعلم، والتكاملات، وخارطة الطريق، والأدلة.	تقنيات الموارد البشرية، أو استشارات التقييم، أو المنتجات، أو أنظمة المواهب.	مدير منتجات، أو مؤسس، أو قائد منصة تقييم.
حوكمة التقييم أو ضمان الذكاء الاصطناعي	السياسة، والمخزون، ومراجعة المخاطر، والتدقيق، وتحدي المؤشرين، وحوكمة النماذج.	مخاطر الموارد البشرية، أو الخصوصية، أو الامتثال، أو التقييم، أو الذكاء الاصطناعي المسؤول.	قائد حوكمة عالمي، أو قائد ذكاء اصطناعي مسؤول، أو تنفيذي مخاطر.
باحث أكاديمي أو تطبيقي	تدريس، وبحث، ونشر، وتطوير منهجي، ومشورة علمية.	دراسة دكتوراه وتدريب بحثي.	أستاذ، أو قائد مركز بحوث، أو مستشار علمي أول.

نقاط قوة قابلة للنقل من الموارد البشرية

نقطة القوة الحالية	كيف تحولها إلى دليل كفاءة في التقييم؟
أطر الكفاءات	أوضح كيف عُرِفَت البُنى وتمييزها وربطها بمؤشرات سلوكية وأساليب وقرارات.
تصميم عملية التوظيف	ابن هندسة اختيار منظمة مع تحليل وظيفة، ومراحل موحدة، وقواعد درجات، ومراجعة عدالة، ومراقبة.
إجراء المقابلات	أنشئ أسئلة مرتبطة بالوظيفة، وتسجيلاً ذا مؤشرات، وتدريباً للمحاورين، ومعايرة، ومراقبة اتفاق المقيمين.
التعلم والتطوير	استخدم النتائج لبناء فرضيات تطويرية قابلة للاختبار، وقدم تغذية راجعة أخلاقية، وقيم تغير السلوك.
استبيانات الموظفين	أثبت تصميم البُنى، وأخذ العينات، وتحليل المقاييس، وإخفاء الهوية، والتكافؤ بين المجموعات، والتقارير، وتقييم الفعل.
تنفيذ تقنيات الموارد البشرية	ارسم البيانات والقرارات، وحدد المتطلبات، واختبر التكاملات، وحوكم الإصدارات، والوصول، والخصوصية، وإمكانية الوصول، وأدلة المورد.
الشراكة مع الأعمال	حوّل طلب صاحب العلاقة إلى قرار ومخاطر وأدلة وميثاق تنفيذ، وتحذّر الافتراضات الضعيفة بصورة بناءة.

أسئلة لمقابلة استكشافية

- ما نسبة الدور المخصصة لأصحاب العلاقة، وتسليم المشروعات، والإحصاء، والكتابة، والإدارة، وعمل المنتج؟
- ما القرارات التي يستطيع الدور اتخاذها مستقلاً، وأيهما يحتاج إلى أخصائي نفسي أو سيكمترى أو محام أو منتدى حوكمة؟
- ما المخرجات الفنية المتوقعة في السنة الأولى؟
- ما البرمجيات والمعايير وأساليب البحث والمؤهلات المستخدمة؟
- كيف تتم مراجعة العمل والإشراف عليه؟
- ما الذي يميز الممارسين الأقوياء عن من يعرفون المصطلحات فقط؟
- ما أدلة المحفظة الأكثر وزناً في التوظيف؟
- ما التوترات الأخلاقية أو التجارية الشائعة، وكيف تتعامل معها المؤسسة؟

ثلاث استراتيجيات للانتقال

الاستراتيجية	كيف تعمل	الأنسب لها
التعمق داخل الدور الحالي	امتلك مسؤولية المقابلات المنظمة، أو مراجعة الموردين، أو قياس الاستبيانات، أو حوكمة التقييم، أو مشروع تحقق.	متخصصو الموارد البشرية الذين يستطيعون الوصول إلى مشروعات مناسبة وإشراف خبير.
الانتقال إلى مزود تقييم أو شركة استشارات	اكتساب خبرة عبر أدوات وعملاء وأساليب وتسليمات متعددة تحت الإشراف.	من يتعلمون بسرعة من المشروعات المتنوعة والعمل مع العملاء.
الدراسة الرسمية والبحث	إكمال دراسة عليا في علم النفس المهني أو I-O أو السيكمتريا أو الأساليب الكمية أو القياس.	من يستهدفون أدوار الأخصائي النفسي أو السيكمترى أو البحث أو التحقق المتقدم.

كيف قد يبدو أسبوع العمل؟

قد يجمع دور الممارس بين ورشة تحليل وظيفة، والتحضير للتفسير، وجلستي تغذية راجعة، واجتماع مزود، ومعايرة مقيمين، وتقرير مراقبة. وقد يقضي السيكمترى وقتاً أكبر في...

...تنظيف بيانات التجريب، واختبار نماذج عاملية أو IRT، والتحقيق في سلوك المفردات، وتوثيق التسجيل، والاجتماع مع فرق المنتج، ومراجعة دراسة. وقد يراجع قائد الحوكمة حالات استخدام مقترحة، ويرأس منتدى مخاطر، ويفحص أدلة تدقيق، ويستجيب لحادث، ويتفاوض على بند مع مورّد، ويقدم إحاطة للقيادات.

العمل نادراً ما يكون «اختبار الناس» طوال اليوم. بل هو تصميم أدلة، وشرح عدم اليقين، وتحسين الأنظمة، وتنسيق الخبرات، وحماية الأشخاص المتأثرين بالقرارات. استراتيجية insyt2 للمسيرة المهنية: اجعل الانتقال مجاوراً لما تقوم به الآن خذ جزءاً من عملك الحالي في الموارد البشرية وارفعه إلى معيار سيكومتري أعلى يمكن إثباته. استخدم هذا الدليل لكسب المستوى التالي من المسؤولية.

22

الفصل

نموذج كفاءة ممارس التقييم

ابن ملفاً على شكل حرف T: اتساع في كفاءات دورة الحياة مع عمق في تخصص أو اثنين.

ثمانية مجالات للكفاءة

المجال	تأسيسي	ممارس	متقدم أو متخصص
علم القياس	يشرح البنى، والثبات، والصدق، والمعايير، والخطأ.	يقيم الأدلة الفنية ويطبقها على استخدام محدد.	يصمم وينتقد نماذج قياس معقدة وحجج تحقق.
الإحصاء والبحث	يفهم التوزيعات، والارتباطات، وأخذ العينات، وعدم اليقين.	ينفذ ويفسر تحليلات أساسية للثبات والصدق والعدالة والنتائج.	يقود دراسات متعددة المتغيرات أو طولية أو IRT أو تكافؤ أو سببية أو تعلم آلي.
تحليل العمل والبنى	يفهم معلومات الوظيفة والمؤشرات السلوكية.	يدير ورش تحليل الوظيفة وينشئ المخططات.	يبني تصنيفات قابلة للتوسع، ونماذج معيار، واستراتيجيات تحقق عبر الأدوار.
تصميم التقييم وتنفيذه	يتبع الإدارة والتسجيل المعياريين.	يصمم مقابلات منظمة، وتمارين، واستبيانات، وتنفيذات مضبوطة.	يقود بنوك مفردات، وأنظمة متعددة الأساليب، وتقييمات تكيفية أو خوارزمية.
العدالة والأخلاقيات والقانون	يتعرف على قضايا النطاق والخصوصية والإعاقة والتمييز.	يدمج التيسيرات، وتحليل العدالة، ومعلومات المشاركين، والتصعيد في المشروعات.	يقود تقييم الأثر، والضمان المستقل، والسياسات عبر الولايات القضائية.
البيانات والتكنولوجيا	يستخدم منصات آمنة ويفهم تدفقات البيانات الأساسية.	يحدد التكاملات، وضبط جودة الدرجات، والوصول، والإصدارات، ولوحات المتابعة.	يحوكم النماذج وواجهات API والأمن والانحراف والبنية والذكاء الاصطناعي المسؤول.
التواصل والتغذية الراجعة	يشرح الدرجات والقيود الأساسية.	يكتب التقارير، وينصح القرارات، ويدرب المستخدمين، ويقدم التغذية الراجعة.	يؤثر في التنفيذيين، ويتعامل مع الأدلة المتنازع عليها، ويترجم عدم اليقين المعقد.
الاستشارات والتنفيذ	يخطط المهام ويدير أصحاب العلاقة.	يحدد نطاق المشروعات، ويدير المخاطر والموردين والميزانيات والتغيير.	يبني ممارسات ومنتجات ونماذج تجارية وحوكمة وقيادة فكرية.

أربعة مستويات لأدلة الكفاءة

تكون الكفاءة أقوى عندما تدعمها أدلة تتجاوز حضور الدورات:

1. المعرفة: تستطيع شرح المفهوم بدقة.
2. التطبيق: تستطيع استخدامه في مهمة محددة النطاق.
3. ممارسة تحت الإشراف: راقب شخص مؤهل عملك وراجع.
4. أثر مستقل: تستطيع قيادة المهمة، وإدارة عدم اليقين، وإنتاج نتائج تحت حوكمة مناسبة.

احتفظ بسجل للتطوير المهني. لكل حالة، سجل السياق، والقرار، ودورك، والأساليب، والإشراف، والحكم الرئيسي، وعدم اليقين، والتغذية الراجعة، والمراجعة، والتعلم. لا تسجل محتوى شخصياً أو سرياً.

مبدأ insyt2 للتعلم: الشهادات مدخلات وليست دليلاً على الكفاءة

قِسْ تقدمك بجودة الأسئلة التي تطرحها، والأدلة التي تستطيع فحصها، والأخطاء التي تمنعها، والعمل الذي تستطيع أدائه بأمان تحت الإشراف المناسب. تقييم ذاتي للجاهزية المهنية قيّم كل عبارة من 1 (ليس بعد) إلى 5 (واثق ولدي دليل). هذه قائمة تأمل وليست أداة سيكمترية متحققاً منها. لكل مجال، سجل دليلاً واحداً على كفاءتك وإجراءً تالياً واحداً.

المجال	عبارة للتأمل	التقييم 5-1
القياس	أستطيع شرح لماذا يتعلق الصدق بتفسير واستخدام مقصودين، لا بالاختيار في حد ذاته.	
القياس	أستطيع تقييم مدى ملائمة مجموعة معيارية وشرح فاصل الثقة.	
القياس	أستطيع التمييز بين أدلة الثبات المناسبة للتقرير الذاتي، واختبارات القدرة، والتمارين التي يقيمها ملاحظون.	
البحث	أستطيع صياغة سؤال تحقق وتحديد تصميم ومعيار مناسبين.	
البحث	أستطيع تفسير حجم الأثر، وعدم اليقين، والبيانات المفقودة، وتقييد المدى، والتحقق المتقاطع على مستوى عملي.	
البحث	أستطيع تحليل نتائج المجموعات الفرعية أو مراجعتها بصورة ذات معنى من دون الاعتماد على إحصائية عدالة واحدة.	
تحليل العمل	أستطيع إدارة عملية منظمة لتحليل الوظيفة أو الحوادث الحرجة.	
تحليل العمل	أستطيع تحويل أدلة الوظيفة إلى تعريفات للبنى ومخطط تقييم.	
تحليل العمل	أستطيع تحدي تسمية غامضة مثل «الإمكانات» أو «الرضا» واستبدالها بأدلة قابلة للملاحظة.	
التصميم والتنفيذ	أستطيع تصميم أو مراجعة جودة الأسئلة والمفردات والتمارين والمؤشرات وبروتوكولات الإدارة.	
التصميم والتنفيذ	أستطيع تحديد التسجيل، والبيانات المفقودة، وإعادة الاختبار، والحوادث، وقواعد القرار قبل النشر.	
التصميم والتنفيذ	أستطيع التخطيط لتجريب وبوابة إطلاق تختبر الأشخاص والعملية والتكنولوجيا والأدلة.	
العدالة والأخلاقيات	أستطيع التمييز بين فرق المجموعة، والأثر السلبي، وتحيز القياس، والتحيز التنبؤي، والعدالة.	
العدالة والأخلاقيات	أستطيع تصميم مسار للتيسيرات ومعرفة متى أحتاج إلى متخصص.	
العدالة والأخلاقيات	أستطيع تحديد حدود نطاق، وتعارض المصالح، ومخاطر الخصوصية، وحقوق المشاركين.	
البيانات والتكنولوجيا	أستطيع تتبع الاستجابات الخام عبر التسجيل والتقارير والقرارات والتخزين والوصول والاحتفاظ والحذف.	
البيانات والتكنولوجيا	أستطيع طرح أسئلة ذات معنى على الموزد بشأن مصدر البيانات والنماذج والإصدارات والانحراف والأمن والتدقيق.	
البيانات والتكنولوجيا	أستطيع اختبار مخرجات التسجيل والتعرف على الحالات التي تخلق فيها خاصية ذكاء اصطناعي مطلب تحقق أو حوكمة جديداً.	
التواصل	أستطيع شرح الدرجات وعدم اليقين والقيود لمدير من دون مصطلحات زائدة أو مبالغة.	
التواصل	أستطيع كتابة تقرير محدد الحدود وإجراء حوار تغذية راجعة بناء.	
التواصل	أستطيع تحدي صاحب مصلحة كبير باحترام عندما يكون الاستخدام المطلوب غير مدعوم.	
الاستشارات	أستطيع تحديد نطاق مشروع من حيث الغرض والنتائج والأدوار والمخاطر والأدلة والخطة والميزانية.	
الاستشارات	أستطيع مقارنة الموزدين باستخدام أدلة موزونة وبوابات حد أدنى.	
الاستشارات	أستطيع تعريف محفزات المراقبة والإجراءات التصحيحية واستراتيجية الخروج.	

تفسير التأمل الذاتي

النمط | دلالة التطوير في الغالب 2-1 | ابدأ بالأساسيات وبعمل تشغيلي محدد النطاق تحت الإشراف. لا تتعجل التفسير المستقل أو تصميم الاختبارات. في الغالب 3 | لديك ثقافة عملية جيدة. ابن عمقاً عبر مشروعات تحت الإشراف تنتج وثائق فنية وأدلة نتائج. في الغالب 4 | قد تكون جاهزاً لدور تخصصي ضمن نطاق واضح. عزز المراجعة المستقلة، والأساليب المتقدمة، والأثر في محفظتك. في الغالب 5 | تحقق من أن التقديرات تعكس أدلة حالية قابلة للملاحظة عبر سياقات مختلفة. فكر في تخصص متقدم، أو قيادة، أو تدريس، أو ضمان. ملف غير متوازن | أمر طبيعي وغالباً مفيد. اختر دوراً مستهدفاً، وعمق المجالات الحرجة له، وأنشئ علاقات إحالة لبقية المجالات.

الفصل

23

المؤهلات والتعلم المتعمد

أربع طبقات للتطوير

الطبقة	الغرض	أمثلة
التعليم التأسيسي	علم النفس، والقياس، والإحصاء، والبحث، والعمل، والمؤسسات.	مقررات جامعية، وشهادات دراسات عليا، وكتب منهجية قوية، ودورات أساليب البحث.
مؤهل مهني أو مؤهل استخدام اختبارات	دليل على أن الممارس يستطيع اختبار أداة وإدارتها وتسجيلها وتفسيرها والتواصل بشأنها ضمن نطاق محدد.	برامج معترف بها لاستخدام الاختبارات المهنية، أو مؤهلات تحت الإشراف، أو تسجيلات مهنية محلية.
تدريب خاص بالأداة	الكفاءة والصلاحيات لاستخدام أداة أو منصة مقيدة لدى ناشر.	دورات القدرة، والشخصية، و360، ومراكز التقييم، أو اعتماد التقارير.
تطبيق تحت الإشراف وتطوير مهني مستمر	تحويل المعرفة إلى ممارسة آمنة والبقاء على اطلاع.	حالات مراجعة، ومعايرة، ومشروعات تحقق، واستشارة أقران، ومؤتمرات، وقراءة، وتدريس، وتدقيق.

اختر التعليم وفق الوجهة المهنية

الدور المستهدف	التعلم ذو الأولوية
مستخدم اختبار مؤهل أو ممارس تغذية راجعة	استخدام الاختبارات المهنية، وتفسير الدرجات، والأخلاقيات، والتغذية الراجعة، والإعاقة والاستخدام العادل، إلى جانب تدريب الأداة.
أخصائي تقييم أو استشاري	علم النفس المهني أو I-O، وتحليل الوظيفة، والاختيار، والأساليب المنظمة، والتحقق، والإحصاء، والعمل مع المشروعات والعملاء.
سيكمتر	نظرية الاختبار الكلاسيكية، والتحليل العاملي، وIRT، ونظرية القابلية للتعميم، والربط، وDIF، والتكافؤ، والبرمجة، وتصميم البحث، والإحصاء المتقدم.
محلل أفراد أو عالم اختبار	الإحصاء، والتصاميم التجريبية والرصدية، وقياس المعايير، والنمذجة التنبؤية، والعدالة، وهندسة البيانات، والتصور.
مدير منتج أو مؤسس	علم التقييم مع اكتشاف المنتج، وتجربة المستخدم UX، والبرمجيات، وواجهات API، وحوكمة البيانات، والأمن، واستراتيجية الأدلة، والتسويق التجاري، ونجاح العميل.
قائد حوكمة أو ذكاء اصطناعي مسؤول	المعايير، وثقافة التحقق، وتقييم الأثر، والخصوصية، والمساواة، وإمكانية الوصول، والأمن، ومخاطر النماذج، والتدقيق، والسياسة.
أخصائي نفسي مهني أو I-O	تعليم معتمد حسب الولاية، وممارسة تحت الإشراف، وأخلاقيات، وتسجيل، وتخصص مختار.

أمثلة لمسارات مهنية

في المملكة المتحدة يدعم مركز الاختبارات النفسية التابع للجمعية البريطانية لعلم النفس BPS مؤهلات استخدام الاختبارات المهنية، ويحتفظ بسجل مؤهلات استخدام الاختبارات، بما في ذلك مسارات القدرة المهنية والشخصية. تستخدم دول وهيئات مهنية أخرى أنظمة مختلفة. وقد يقيد الناشرون الأدوات حسب مستوى التأهيل. تحقق دائماً من متطلبات الدخول الحالية، والاعتراف، والتجديد، والنطاق، وما إذا كانت الدورة تمنح صلاحية استخدام أداة بعينها.

يمكن للدراسة العليا في علم النفس المهني أو الصناعي-التنظيمي أن توفر تدريباً علمياً وتطبيقياً واسعاً. وتوفر برامج السيكمترية، وعلم النفس الكمي، والقياس التربوي، والإحصاء، وعلم البيانات، وأساليب البحث خبرة أعمق في القياس.

يعتمد المسار الأقوى على العمل الذي تريد القيام به وعلى التنظيم المهني في المكان الذي ستمارس فيه. لا تمنح دورة قصيرة لقباً مهنيّاً محمياً ولا كفاءة واسعة في تصميم التقييم. كيف تقيّم دورة أو مؤهلاً؟

- هل المنهج متوافق مع معايير معترف بها ومع المهام التي تريد أداؤها؟
- هل يتضمن ممارسة مقيمة، لا مجرد فيديوهات واختبارات قصيرة؟
- من يدرّس ويشرف، وما خبرته الحالية؟
- هل ستفحص أدلة فنية حقيقية وتنتج أعمال تفسير أو تصميم؟
- ما النطاق أو اللقب أو السجل أو صلاحية الوصول للأدوات التي يمنحها الإكمال تحديداً؟
- هل تُقيّم الكفاءة بصورة مستقلة؟ وهل يمكن عدم منح المؤهل لمن لا يحقق المستوى؟
- ما نظام التطوير المستمر، والتجديد، والأخلاقيات، والشكاوى، أو الإشراف؟
- هل سيكون المؤهل مفهوماً ومعترفاً به في سوق العمل الذي تستهدفه؟

مجالات الدراسة الأساسية

المجال	الموضوعات الدنيا
القياس	البنى، ونظرية الاختبار الكلاسيكية، والثبات، وSEM، والصدق، والمقاييس، والمعايير، وتحليل المفردات، والنماذج العالمية، وأساسيات IRT.
تقييم الأفراد	تحليل الوظيفة، والمتنبئات، والمقاييل المنظمة، وعينات العمل، وSJT، والشخصية، ومراكز التقييم، وقواعد القرار، والمنفعة.
الإحصاء والبحث	أخذ العينات، والتوزيعات، وحجم الأثر، وفترات الثقة، والارتباط والانحدار، والبيانات المفقودة، والقوة الإحصائية، والتحقق المتقاطع، والتصميم الطولي.
العدالة وإمكانية الوصول	التحيز، وDIF، والتكافؤ، والتنبيؤ حسب المجموعات، والأثر السلبي، والتيسيرات، والتصميم الشامل، والتكيف عبر الثقافات.
الأخلاقيات والقانون	الكفاءة، ومعلومات المشارك، والسرية، والإعاقة، والتمييز، وحماية البيانات، والمعايير، والقرارات المؤتمتة.
التكنولوجيا	المنصات، ورسم البيانات، وضبط جودة الدرجات، وواجهات API، والأمن، وهندسة الخصوصية، وأداء النماذج، وقابلية التفسير، والانحراف، والإشراف البشري.
الممارسة المهنية	التغذية الراجعة، وكتابة التقارير، والتيسير، والاستشارات، وتحدي أصحاب العلاقة، والحوكمة، والمهارات التجارية، وتسليم المشروعات.

البرمجيات وأدوات التحليل

يُعد Excel مفيداً لفحص الدرجات، والتحليل الوصفي، ونماذج القرار، ولوحات المتابعة، والنماذج الأولية السهلة، لكنه يحتاج إلى صيغ منضبطة، وضبط إصدارات، وحالات اختبار، وحماية. يدعم SPSS التحليل الإحصائي القياسي بواجهة مألوفة. ويوفر R وPython تحليلاً قابلاً لإعادة الإنتاج والتوسع للنماذج العالمية، وIRT، والمحاكاة، والعدالة، والتصوير، والأتمتة. ويساعد SQL على العمل مع البيانات التشغيلية. لا تحل أداة محل جودة التصميم أو البيانات أو التفسير.

تطور على مراحل:

1. تعلم حساب وشرح المتوسطات، والانحرافات المعيارية، ودرجات z وT، وSEM، وفترات الثقة، والارتباطات، ومعدلات الاختيار، والمركبات البسيطة.
2. تعلم تنظيف البيانات، وقرارات البيانات المفقودة، والشفرة القابلة لإعادة الإنتاج، وضبط الإصدارات، والتصور الأساسي.
3. أضف الانحدار، والثبات، والتحليل العاملي، وتحليل المجموعات الفرعية، وتقارير التحقق.
4. أضف الأساليب المتقدمة فقط عندما يتطلبها الدور ويمكنك الحصول على تدريب ومراجعة عاليي الجودة.

حلقة تعلم متعمد

1. اختر دوراً مستهدفاً واحداً وفجوتين ذات أولوية للربع القادم. 2. ادرس مفهوماً من مصدر موثوق. 3. طبقه على مخرج صغير حقيقي أو محاكى. 4. اطلب من شخص كفؤ مراجعة العمل والمنطق معاً. 5. عدّل المخرج وسجل ما تغير. 6. اشرح المفهوم لجمهور موارد بشرية غير تقني. 7. استخدمه في مشروع تحت الإشراف بعواقب متناسبة مع كفاءتك. 8. أضف المخرج النهائي، والتأمل، ودليل الأثر إلى المحفظة. معيار insyt2 للكفاءة: اجعل حدودك صريحة يستطيع الممارس الموثوق تحديد نطاق كفاءته، والأدلة التي تدعمه، والإشراف الذي يستخدمه، والنقطة التي يلجأ فيها إلى الإحالة.

24

الفصل

خارطة انتقال لمدة اثني عشر شهراً

ابن الأساسيات، ومخرجات تحت الإشراف، وتخصصاً، ومحفظة مهنية مرئية خلال عام واحد.

الأشهر 1-3: الأساسيات والتوجيه

الشهر	التركيز	المخرجات
1	الهدف المهني وخط الأساس	اختر دورين مستهدفين، وأكمل التأمل في الكفاءة، وأجر مقابلات مع ثلاثة ممارسين، واكتب خطة تطوير.
2	أساسيات القياس	ادرس البنى والثبات والصدق والمعايير والعدالة؛ واكتب مراجعة فنية من صفحتين لأداة مألوفة.
3	تحليل الوظيفة والدليل المنظم	نفذ تمرين حوادث حرجة تحت الإشراف وأنشئ تعريفات للبنى ومخطط تقييم.

خلال هذه المرحلة لا تحاول بناء اختبار مملوك كامل. تعلم قراءة الأدلة الفنية، وطرح أسئلة أفضل، وتوثيق السلسلة من العمل إلى البنية إلى الدليل. انضم إلى مجتمعات مهنية ذات صلة، وحدد مشرفاً أو مجموعة مراجعة أقران، وأنشئ بيئة محمية لمحفظتك لا تتضمن بيانات سرية أو مفردات مقيدة.

الأشهر 4-6: استخدام الاختبارات والتصميم التطبيقي

الشهر	التركيز	المخرجات
4	كفاءة مستخدم الاختبار	أكمل برنامجاً مناسباً لاستخدام الاختبارات المهنية أو دورة ناشر، وتمرن على التفسير والتغذية الراجعة تحت الإشراف.
5	تصميم أسلوب منظم	ابن دليل مقابلة أو محاكاة يتضمن أسئلة ومؤشرات وتدريب مقيمين وضوابط جودة التسجيل.
6	البيانات والتحليل	حلل مجموعة بيانات مجهلة أو محاكاة للثبات، وتوزيعات الدرجات، والعلاقات بالمعيار، ونتائج المجموعات الفرعية، واذكر القيود.

اختر مشروعات تظهر منطقك لا شرائح جميلة فقط. أدرج المخطط، والبدائل التي نوقشت، والافتراضات، وتعليقات المراجع، والتعديلات، وخطة المراقبة. اطلب من المراجعين تحدي صلة البنى، والتفسير الإحصائي، والعدالة، والتواصل.

الأشهر 7-9: التحقق والحوكمة والتخصص

الشهر	التركيز	المخرجات
7	تصميم التحقق	اكتب خطة دراسة تنبؤية أو متزامنة تتضمن المعايير والعينة والفرضيات والتحليلات والعدالة ومحفزات القرار.
8	المورد والحوكمة	أكمل بطاقة عناية واجبة وصمم مخزون تقييمات وفئة مخاطر وبوابة إطلاق.
9	دفعة تخصصية	اختر التغذية الراجعة، أو تصميم الاختبار، أو الاستبيانات، أو السيكومتريا، أو التحليلات، أو ضمان الذكاء الاصطناعي، أو المنتج، وأكمل مشروعاً أعمق.

الأشهر 10-12: الدليل المهني ودخول السوق

الشهر	التركيز	المخرجات
10	الأثر والتواصل	قدم حالة لقادة الموارد البشرية: المشكلة، والدليل، وتصميم القرار، والمخاطر، والنتيجة، والدروس؛ وأنشئ ملخصاً تنفيذياً.
11	المحفظة والتموضع	انشر أو شارك بصورة خاصة ثلاثة مخرجات منقحة وخالية من الحساسية؛ وحذث السيرة الذاتية، والملف المهني، وقصة الدور، والطلبات المستهدفة.
12	مشروع ختامي والقرار التالي	أكمل مشروعاً أو تدقيقاً كاملاً تحت الإشراف؛ وراجع أدلة الكفاءة واختر المؤهل أو الدور التالي.

أول تسعين يوماً بالتفصيل

1. اكتب هدفاً من جملة واحدة: «خلال اثني عشر شهراً أريد أن أكون موثقاً للتقدم إلى [الدور] لتنفيذ [المخرجات] في [السوق]».
2. اختر عملية موارد بشرية حية يمكنك تحسينها من دون تجاوز نطاقك، مثل تنظيم مقابلة أو مراجعة دليل فني لمورد.
3. اقرأ معياراً مهنيّاً واحداً ودليلاً فنياً واحداً بصورة نشطة، ولخص الادعاءات والأدلة والقيود والأسئلة.
4. أنشئ ملف حسابات صغيراً أو شفرة لحساب Z و T و SEM وفترات الثقة ومعدلات الاختيار ومركب توضيحي.
5. راقب أو شارك في تيسير جلستي تغذية راجعة أو تحليل وظيفة بإذن وإشراف.
6. أنشئ مسرداً بكلماتك وشرح خمسة مفاهيم لزميل غير تقني.
7. ابحث عن مراجع واتفق معه على ما سيفحصه: منطق البنى، والإحصاء، والعدالة، ولغة التقارير، والنطاق.
8. ابدأ سجل حالات يوثق الغرض، والأدلة، والقرارات، وعدم اليقين، والإشراف، والتعلم.

إيقاع تعلم أسبوعي

النشاط	الإيقاع المقترح
دراسة فنية	جلستان مركزتان أسبوعياً باستخدام أمثلة محلولة بدلاً من القراءة السلبية.
مخرج تطبيقي	مخرج صغير كل أسبوعين: مخطط، أو دليل تسجيل، أو تحليل، أو تقرير، أو جزء من تدقيق.
إشراف أو مراجعة أقران	شهرياً على الأقل، وبشكل أكثر تكراراً للتفسير الفردي أو التصميم عالي المخاطر.
قراءة مهنية	بحث واحد، أو قسم من معيار، أو إرشاد موثوق كل أسبوع.
تأمل	بعد كل مشروع: ما الاستنتاج؟ ما الدليل الذي دعمه؟ ما الذي قد يكون خاطئاً؟ وما الذي تغير؟
تعلم السوق	محادثة مع ممارس أو تحليل دور واحد شهرياً لإبقاء الخطة مرتبطة بالسوق.

علامات التقدم

- تطلب الدليل الفني قبل طلب العرض التجريبي.
- تصف ما لا تعنيه الدرجة بثقة مساوية لما تعنيه.
- تستطيع شرح فاصل الثقة، والمجموعة المعيارية، وقيد التحقق لمدير.

• تحسن قاعدة القرار بدلاً من مجرد إضافة مرحلة تقييم.

• تكتشف مخاطرة تتعلق بإمكانية الوصول أو البيانات قبل أن يواجهها المرشح. • تغير توصيتك عندما تتغير الأدلة. • يستطيع ممارس كفؤ آخر مراجعة عملك وإعادة إنتاجه. اختبار تقدم insyt2: ما الذي تستطيع الآن القيام به بأمان؟ السؤال الحقيقي ليس عدد الدورات التي أكملتها، بل ما عمل التقييم الذي تستطيع الآن تنفيذه بأمان، وشرحه بوضوح، والدفاع عنه بالأدلة.

الفصل

25

المحفظة المهنية ومسارات التوظيف والاستشارات

سنة مشروعات للمحفظة

المشروع	ما الذي ينبغي أن يثبتته؟
تحليل وظيفة ومخطط	الحوادث الحرجة، وتعريفات البنى، والمؤشرات السلوكية، وتغطية المحتوى، واختيار الأسلوب، والتحقق مع أصحاب العلاقة.
مقابلة منظمة أو محاكاة	الأسئلة أو التمرين، ومؤشرات التسجيل، ودليل المقيم، وخطة التدريب، والثبات، وتصميم المراقبة.
مراجعة فنية لاختبار	الغرض، والبنية، والمعايير، والثبات، والصدق، والعدالة، وإمكانية الوصول، والأمن، والقيود، والتوصية.
تحليل تحقق	سؤال البحث، وجودة البيانات، والأدلة الوصفية، والثبات، والعلاقة بالمعيار، وتحليل المجموعات الفرعية، وعدم اليقين، والقيود.
حزمة مرشح وحوكمة	الإشعار، والاستعداد، والتيسيرات، والحوادث، والمراجعة، والخصوصية، وضبط جودة التسجيل، والمخزون، وبوابة الإطلاق.
دراسة حالة تنفيذية	مشكلة الأعمال، وهندسة القرار، والبدائل، والأدلة، والتنفيذ، والنتائج، والمخاطر، والدروس بلغة واضحة.

قواعد المحفظة

- استخدم بيانات محاكاة أو عامة أو مجهلة بصورة صحيحة. لا تكشف محتوى مرشح أو موظف أو عميل أو مفردات مقيدة.
- حدد دورك وإشرافك ومساهمة الآخرين.
- اعرض مسودات مختارة وتعديلات حتى يرى المراجعون كيف تستجيب للتحدي.
- أدرج الافتراضات والقيود والقرارات التي لم تتخذ، لا النتائج الإيجابية فقط.
- اشرح الجمهور والقرار. التحليل الصحيح فنياً الذي لا يستطيع توجيه الممارسة غير مكتمل.
- لا تختلق نتائج تحقق ولا توج بأن أداة عرض تجريبي متحقق منها تشغيلياً.
- عند مراجعة أداة مملوكة، ناقش فئات الأدلة والاستنتاجات من دون إعادة إنتاج مفردات محمية أو محتوى الدليل الفني.

التموضع في السيرة الذاتية والمقابلة

استبدل الادعاءات العامة مثل «خبرة في السيكمترية» بأدلة. مثال: «صممت مقابلة منظمة لسنة أدوار انطلاقاً من 42 حادثة حرجة؛ دربت 18 محاوراً؛ وطبقت تسجيلاً ذا مؤشرات، ومعايرة، ومراقبة للاستثناءات». وفي المقابلة كن مستعداً لشرح قرار غيرته بعد ظهور أدلة، وقيد أفصحت عنه، وتعارض تعاملت معه، وموقف أحلت فيه العمل إلى خبير أكثر تأهيلاً.

ادعاء ضعيف	بديل قائم على الأدلة
«معتمد في عدة اختبارات».	«مؤهل لاستخدام أدوات مهنية محددة؛ أكملت تفسيرات وجلسات تغذية راجعة تحت الإشراف؛ وأحافظ على التطوير المهني وضوابط النطاق».
«بنيت تقييماً سيكومترياً».	«عرّفت البنية والمخطط، وكتبت وراجعت المفردات، وجرّبتها، وحللت البنية الأولية والثبات، ووثقت التحقق الذي ما زال مطلوباً».
«استخدمت الذكاء الاصطناعي لتحسين التوظيف».	«قدت مراجعة أثر وأدلة لأداة فرز بالذكاء الاصطناعي، وأزلت خصائص غير مدعومة، وأضفت مراجعة بشرية ومراقبة انحراف، ووثقت المخاطر المتبقية».
«متخصص موارد بشرية قائم على البيانات».	«صممت دراسة معيار، وطابقت التسجيل، وأبلغت عن عدم اليقين ونتائج المجموعات الفرعية، وغيرت قاعدة القرار بعد التحقق المتقاطع».

الاستشارات المستقلة

تتطلب الاستشارات أكثر من معرفة التقييم. عرّف الخدمات، والعملاء المستهدفين، وحدود الكفاءة، والتأمين المهني، والهيكّل القانوني، والعقود، والسرية، وأدوار معالجة البيانات، وضمان الجودة، والإشراف، وشبكات الإحالة. حوّل الخدمات المحددة إلى عروض واضحة قبل الوعد بتطوير اختبارات مخصصة بالكامل. ومن الأمثلة: تصميم المقابلات المنظمة، وتدقيق التقييم، والعناية الواجبة بالموردين، وتدريب المقيمين، ومراجعة جودة الاستبيانات، وتيسير التغذية الراجعة.

- استخدم بيانات عمل مكتوبة تحدد الغرض، والمخرجات، والأدلة، والمسؤوليات، والاستثناءات، ومعايير القبول.
- افصل أتعاب المراجعة المستقلة عن مبيعات الأدوات أو أفصح عن العلاقة.
- احمِ محتوى الاختبارات وبيانات العملاء عبر أنظمة معتمدة، وضوابط وصول، واحتفاظ، وإتلاف آمن.
- استخدم مراجعة أقران للتقارير الفنية والتوصيات عالية المخاطر.
- حافظ على قوالب، وضبط إصدارات، وسجلات حالات، وتطوير مهني، وعملية شكاوى.
- سّعّر التحليل والتوثيق والمراجعة والحوكمة، لا وقت الاجتماعات فقط.
- ارفض العمل خارج كفاءتك حتى لو كان العميل مستعداً للدفع.

بناء منتج تقييم أو منصة SaaS

عمل SaaS للتقييم هو في الوقت نفسه برنامج قياس، ومنتج برمجي، وخدمة معالجة بيانات، ونظام أمن، ومنظومة ممارسة مهنية، ومؤسسة تجارية. الحد الأدنى للمنتج القابل للحياة ليس أصغر استبيان ينتج تقريراً، بل أصغر نظام يستطيع تقديم وعد محدود ومدعوم بالأدلة بصورة آمنة.

مسار عمل المنتج	الحد الأدنى من الأسئلة التي تثبت المصادقية
العلم	ما البنية والاستخدام اللذان تدعمهما الأداة؟ ما الأدلة الموجودة، وما الأدلة المخطط لها، وما الادعاءات المحظورة؟
المحتوى	كيف تُحدّد مواصفات المفردات أو التمارين، وتراجع، وتُجرب، وتؤمّن، وتُحدّث، وتُدار إصداراتها؟
التسجيل	هل يمكن إعادة إنتاج كل درجة؟ كيف تُضبط البيانات المفقودة، ودرجة الثقة، والمعايير، والفئات، والدرجات المركبة، والتحديثات؟
تجربة المستخدم	هل يستطيع المشاركون فهم العملية والوصول إليها وإكمالها وإيقافها مؤقتاً والحصول على الدعم وطلب التيسيرات ومراجعة القرارات الجوهرية؟
البنية التقنية والأمن	هل صُممت بنية تعدد المستأجرين، والتشفير، والهوية، والوصول، والسجلات، والنسخ الاحتياطية، والحوادث، وواجهات API، والمعالجون الفرعيون بما يتناسب مع مخاطر التقييم؟
البيانات والخصوصية	من هو المتحكم أو المعالج؟ ما الذي يُجمع، ولماذا، وأين، وكم من الوقت، وما الحقوق، وعمليات النقل، والحذف؟
الاستخدام المهني	أي المستخدمين يحتاجون إلى مؤهل؟ ما التدريب وفحوص الكفاءة المطلوبة؟ وكيف يُمنع سوء الاستخدام؟
عمليات الأدلة	كيف تُدار مجموعات بيانات البحث، وتحليلات العدالة، والانحراف، ودراسات العملاء، والتقارير الفنية؟
الادعاءات التجارية	هل يمكن للتسويق إسناد كل ادعاء إلى دليل حديث مع التعبير الدقيق عن القيود؟
الحوكمة	من يملك صلاحية اعتماد إصدار، أو تغيير نموذج، أو استخدام جديد، أو سوق جديدة، أو تراجع عاجل عن إصدار، أو سحب المنتج؟

استراتيجية مرحلية للمنتج والأدلة

1. المفهوم: مراجعة البنية، والحاجة السوقية، والجدوى الأخلاقية والقانونية، وخطة الأدلة.
2. النموذج البحثي الأولي: المواصفات، ومراجعة الخبراء، والاختبارات المعرفية واختبارات سهولة الاستخدام؛ من دون ادعاء تشغيلي عالي المخاطر.
3. التجربة: عينة مضبوطة، وتحليل المفردات والمقاييس، وأدلة أولية على الثبات والبنية، ونتائج إمكانية الوصول والتشغيل.
4. إطلاق محدود النطاق: استخدامات محددة بوضوح، ومستخدمون مدربون، ومراقبة، ودراسات محلية، ودليل استخدام، وضبط للإصدارات.
5. نطاق متحقق منه: أدلة تراكمية من مصادر متعددة، ومعايير أو مستويات مناسبة، وتحليل للعدالة، وأدلة على القرار، ومراجعة مستقلة.
6. التوسع: كل دور جديد أو لغة أو دولة أو نمط تقديم أو قرار أو خاصية ذكاء اصطناعي يخضع لتقييم قابلية نقل الأدلة والمخاطر.

أخطاء مهنية شائعة

- جمع عدد كبير من شهادات الناشرين من دون فهم أسس القياس.
- وصف كل استبيان بأنه سيكمترية أو كل ارتباط بأنه تحقق.
- تجنب الإحصاء إلى درجة تعجز معها عن تقييم ادعاءات الموزد.
- التركيز على الإحصاء إلى درجة إهمال صلة الأداة بالوظيفة وتجربة المرشح والتنفيذ والأخلاقيات.
- نشر محتوى مقيد أو مواد حالات سرية ضمن المحفظة المهنية.
- الوعد بتقييم «خال من التحيز» أو «موضوعي بالكامل» أو «محايد ثقافياً».
- العمل بلا إشراف أثناء مرحلة تعلم التفسير الفردي أو التصميم عالي المخاطر.
- التعامل مع التقرير الجذاب باعتباره بديلاً عن الدليل الفني.
- بناء منتج قبل تحديد البنية والاستخدام وادعاء الدليل ونموذج الحوكمة.
- السماح للموعد التجاري بأن يصبح هو معيار التحقق.

منظور ختامي

تُبنى المسيرة الناجحة في التقييم السيكومتري على فضول منضبط. اسأل: ماذا تعني الدرجة؟ ماذا تفوت؟ من قد يتضرر منها بصورة غير متناسبة؟ ما الدليل الذي سيغيّر رأيك؟ وكيف ستعرف المؤسسة ما إذا كان النظام يعمل فعلاً؟ يكافئ هذا المجال من يستطيع الاحتفاظ بحقيقتين في الوقت نفسه: القياس يمكن أن يحسن القرارات البشرية، والقياس يظل دائماً غير كامل. والمصادقية المهنية تأتي من إدارة هذا النقص بمسؤولية.

الخطوة التالية
من insyt2

حوّل التعلم إلى مشروع يحمل دليلاً

اختر مشكلة تقييم قريبة من عملك، وعَرّف القرار، وقوِّ سلسلة الأدلة، واحصل على مراجعة خبير، ووثّق القيود.

المعيار المهني
من insyt2

ابن دليلاً على الكفاءة، لا مجرد معرفة بالأدوات.

تعلّم العلم. مارس تحت إشراف. وثّق الأدلة. واعرف متى تحيل الأمر إلى متخصص. التقييم يجب ألا ينتهي بدرجة.

الملاحق

الملاحق

قوالب عملية، وقوائم فحص، وصيغ، وتعريفات، ومصادر مرجعية للعمل المسؤول في التقييم. الملحق العملي A

ورقة الصيغ والتفسير العملية

هذا الملحق مرجع عمل، وليس بديلاً عن الدليل الفني للأداة أو المراجعة الإحصائية. استخدم تعريفات الدرجات، والنموذج، وتقديرات الثبات، والمعايير، والأخطاء المعيارية الخاصة بالأداة المحددة.

الدرجات المعيارية

$$z = (X - M) / SD$$

حيث X الدرجة الخام، وM متوسط المجموعة المرجعية الملائمة، وSD انحرافها المعياري.

$$T = 50 + 10z$$

تكون درجة T ذات متوسط مرجعي 50 وانحراف معياري 10 عندما يُطبق التحويل كما هو محدد.

تشير درجة $z = +1.0$ إلى انحراف معياري واحد فوق المتوسط المرجعي؛ وتمثل $T = 60$ الموقع المعياري نفسه. لا يجعل التوحيد التوزيع طبيعياً، ولا يزيل خطأ القياس، ولا يجعل مجموعة معيارية غير ملائمة مناسبة.

الخطأ المعياري للقياس وفترات الدرجة

$$SEM = SD \times \sqrt{1 - \text{الثبات}}$$

يجب أن يتوافق تقدير الثبات مع الدرجة المبلّغ عنها، والسكان، وطريقة الإدارة.

فترة تقريبية للدرجة عند 95% = الدرجة المرصودة $\pm 1.96 \times SEM$

استخدم طريقة الثقة المحددة في الدليل متى كانت متاحة؛ وقد تختلف الفترات وفق نظرية الاستجابة للمفردة أو الدقة الشرطية بحسب مستوى الدرجة.

مثال: إذا كان $SD = 10$ والثبات = .84، فإن $SEM = 10 \times \sqrt{(.16)} = 4$. ودرجة مرصودة قدرها 62 تكون الفترة التقريبية البسيطة عند 95% هي 62 ± 7.84 ، أي نحو 54 إلى 70. يوضح ذلك لماذا لا ينبغي التعامل مع الفروق الصغيرة في الترتيب على أنها دقيقة تماماً.

صيغة سبيرمان-براون للتخطيط

الثبات المقدّر بعد تغيير الطول = $[1 + (k - 1) \times r] / (k \times r)$

حيث k نسبة الطول الجديد إلى الطول القديم، و r تقدير الثبات الحالي.

تقدّر الصيغة أثر إضافة محتوى متوازٍ وبجودة مشابهة وفق افتراضات النظرية الكلاسيكية للاختبار. ويمكن أن تؤدي إضافة مفردات ضعيفة أو زائدة إلى زيادة الطول من دون تحسين الصدق أو تغطية المجال.

نتائج الاختيار

معدل الاختيار = عدد المختارين / عدد من خضعوا للتقييم

عرّف السكان والفترة والمرحلة وطريقة التعامل مع المنسحبين بصورة متسقة.

نسبة الأثر السلبي = معدل اختيار المجموعة / أعلى معدل اختيار ذي صلة بين المجموعات

تُعد نسبة أقل من 80 إشارة للمراجعة في سياق الإرشادات الموحدة الأمريكية، وليست حكماً قانونياً كاملاً.

مثال: معدل اختيار المجموعة $A = 50\%$ ، والمجموعة $B = 35\%$. النسبة = $35 / 50 = 70$. افحص المرحلة، وحجم العينة، وصلة الإجراء بالوظيفة، والبدائل، والقياس، والإدارة، وقاعدة القرار.

فرق المتوسط المعياري

$d = (\text{متوسط المجموعة 1} - \text{متوسط المجموعة 2}) / (\text{الانحراف المعياري المجمع})$

يُبين اتجاه الفرق، والمجموعات، والعينة، وعدم اليقين، وما إذا كانت الدرجة تحمل معنى قابلاً للمقارنة بين المجموعات.

فرق المتوسط وصف للبيانات، ولا يثبت بمفرده تحيز المفردة أو التحيز التنبؤي أو التمييز أو العدالة.

مقاييس التصنيف

المقياس	الصفة	السؤال المبسط
الحساسية	الإيجابيات الحقيقية / جميع الإيجابيات الفعلية	من بين من يحققون المعيار، كم شخصاً تم تحديده؟
النوعية	السلبيات الحقيقية / جميع السلبيات الفعلية	من بين من لا يحققون المعيار، كم شخصاً تم رفضه بصورة صحيحة؟
القيمة التنبؤية الإيجابية	الإيجابيات الحقيقية / جميع الحالات المتنبأ بأنها إيجابية	من بين المختارين، كم شخصاً يحقق المعيار؟
القيمة التنبؤية السلبية	السلبيات الحقيقية / جميع الحالات المتنبأ بأنها سلبية	من بين المرفوضين، كم شخصاً لا يحقق المعيار؟
معدل الإيجابيات الكاذبة	الإيجابيات الكاذبة / جميع السلبيات الفعلية	كم مرة يُختار شخص لا يحقق المعيار؟
معدل السلبيات الكاذبة	السلبيات الكاذبة / جميع الإيجابيات الفعلية	كم مرة يُرفض شخص قادر؟

تعتمد القيم التنبؤية بدرجة كبيرة على معدل الأساس وجودة المعيار. وفي التوظيف نادراً ما تكون «الحقيقة» الخاصة بالمعيار مرصودة على نحو كامل.

الارتباط والتباين المفترس

يتراوح ارتباط بيرسون بين -1 و $+1$ ويلخص العلاقة الخطية في شروط محددة. وهو لا يثبت السببية. ويصف مربعه r^2 نسبة التباين المشترك في نموذج بسيط ذي متغيرين، لكن لا ينبغي تحويل ذلك إلى ادعاء بأن التقييم «يفسر الشخص».

أبلغ عن الارتباط المرصود، وفترة الثقة، والمعيار، وتقييد المدى، والبيانات المفقودة، وتصميم الدراسة، والدلالة العملية. قد تكون تقديرات الصدق المصححة مفيدة، لكنها تعتمد على افتراضات؛ اعرض القيم المرصودة والمصححة معاً بدلاً من المصححة وحدها.

درجة مركبة بسيطة

الدرجة المركبة = $w1(z1) + w2(z2) + ... + wk(zk)$

حوّل الدرجات إلى مقياس مشترك يمكن الدفاع عنه، ووثّق الأوزان، والاتجاه، وقواعد البيانات المفقودة، وقواعد العتبات.
مثال توضيحي: $z_{0.40}$ الاستدلال + $z_{0.35}$ المقابلة_المنظمة + $z_{0.25}$ عينة_العمل. هذه الأوزان مثال فقط. تتطلب الأوزان الفعلية تحليل الوظيفة، والنبات، والصدق، والتدخل بين الأساليب، والعدالة، والمنفعة، والتحقق المتقاطع.

أسئلة المنفعة

- يمكن لتحليل المنفعة الكامل أن يشمل الصدق، ونسبة الاختيار، ومعدل الأساس، وعدد المعينين، ومدة البقاء، وفروق قيمة الأداء، والتكاليف، وآثار التنفيذ. وحتى من دون نموذج مالي رسمي أسأل:
- ما خطأ القرار الذي يُفترض أن يقلله التقييم؟
 - كم قراراً سيبتأئ، ولفترة كم؟
 - ما تكلفة الأسلوب على المشاركين والمؤسسة؟
 - هل يضيف دليلاً يتجاوز الأساليب الحالية؟
 - ما تكاليف العدالة وإمكانية الوصول والخصوصية والسمعة المحتملة؟
 - ما مقدار تحسن النتائج الذي سيجعل الاستثمار مجدياً؟
 - كيف سيقيم أثر المنفعة بعد الإطلاق؟
- ملاحظات التفسير المحلي
- الأداة والإصدار ومقياس الدرجات
- المجموعة المرجعية والسكان المستهدفون

أدلة الثبات والخطأ المعياري للقياس

- قاعدة القرار أو العتبة أو نطاق المراجعة
- تحفظات العدالة وإمكانية الوصول
- مالك الاعتماد وتاريخ المراجعة
- ملاحظة عمل: سجّل مصادر الأدلة، والافتراضات، والقيود، والمسؤولين، وتاريخ المراجعة.
- الملحق العملي B
- بطاقة تقييم المورد
- كيفية الاستخدام
1. عزّف الاستخدام المقصود ومستوى المخاطر قبل التواصل مع الموردين.
 2. حدد الأوزان قبل العروض التوضيحية.
 3. استخدم مقياساً من 0 إلى 5 لجودة الدليل، لا لجودة العرض البصري.
 4. ضع متطلبات اجتياز إلزامية للمعايير الحرجة؛ فلا يمكن لمجموع مرتفع أن يعوض فشلاً جوهرياً في الدليل أو الخصوصية أو الأمن.
 5. سجّل فقط الأدلة التي تمت مراجعتها فعلياً، ووثّق ما هو مفقود.
 6. وفق مستوى المخاطر، أشرك الاختصاصات السيكمترية، وإمكانية الوصول، والقانون، والخصوصية، والأمن، والتقنية، والمشتريات، والعمليات، والمستخدمين.

الدرجة	تفسير مستوى الدليل
0	لا دليل، أو دليل متناقض، أو رفض لتقديمه.
1	ادعاء فقط أو مواد تسويقية.
2	دليل محدود، أو عينة ضعيفة، أو صلة غير واضحة بالاستخدام.
3	دليل مقبول مع قيود جوهريّة معلنة.
4	دليل قوي وذو صلة مع شفافية في القيود.
5	دليل قوي ومتكرر ومناسب للاستخدام وخاضع لتحذ أو مراجعة مستقلة.

الدرجة الموزونة = مجموع [(درجة المعيار / 5) × وزن المعيار]. أبقى الحد الأقصى 100 واعرض متطلبات الاجتياز منفصلة.

المعيار	الوزن (%)	متطلب اجتياز؟
صلة الأداة بالوظيفة ووضوح البنية	15	نعم
الصدق للاستخدام المقصود	15	نعم
الثبات والدقة	10	نعم
المعايير أو المستويات ودعم التفسير	8	نعم
العدالة وأدلة المجموعات الفرعية	10	نعم
إمكانية الوصول والتيسيرات	8	نعم
شفافية التسجيل وحوكمة الخوارزميات	8	نعم إذا كان التسجيل مؤتمتاً
الخصوصية والأمن ونزاهة الاختبار	10	نعم
تجربة المرشح والمستخدم الإداري	5	لا
العمليات والدعم والتكامل	4	لا
المراقبة والتدقيق وضبط التغيير	4	نعم
القيمة التجارية والتكلفة الكلية	3	لا

الحد الأدنى من أسئلة طلب العروض (RFP)

الغرض والبنية

- اذكر كل استخدام مدعوم وغير مدعوم، والسكان، واللغات، ومستويات الأدوار، وأنواع القرارات.
- عرّف كل بنية يتم الإبلاغ عنها، واعرض نموذج الأدلة المرتبط بالعمل.
- اشرح ما الذي لا تقيسه الدرجة، وما الاستنتاجات التي يجب ألا يقوم بها المستخدمون.

التطوير والتسجيل

- قدّم مخطط المحتوى، وعملية تطوير المفردات أو التمارين، ومؤهلات الخبراء، وتصميم التجربة، وضوابط المراجعة.
- اشرح التسجيل الخام، والتحويلات، والتعامل مع البيانات المفقودة، ومؤشرات جودة الاستجابة، وفترات الثقة، والمعايير، والفئات، والدرجات المركبة، وقواعد الحدود.
- حدّد الإصدارات الحالية، وكيف يتم الإخطار بالتغيير، وكيف يمكن التراجع عن إصدار عند الحاجة.

الثبات والصدق والعدالة

- قدّم الأدلة لكل درجة يتم الإبلاغ عنها، لا للدرجة الكلية فقط.
- اعرض العينات والتواريخ والوظائف والدول واللغات والمعايير وعدم اليقين والتحقق المتقاطع.
- قدّم تحليلات الأداء التفاضلي للمفردة (DIF)، وتكافؤ القياس، والتنبؤ حسب المجموعات الفرعية، ونتائج الاختيار، وإمكانية الوصول، وتجربة المرشح حيثما كانت ذات صلة.
- افصح عن الأدلة غير المواثبة أو غير الحاسمة وعن حدود المنتج.

التكنولوجيا والذكاء الاصطناعي

- اذكر كل خاصية تؤثر في الدرجة أو التوصية، وشرح صلتها بالوظيفة.
 - اشرح بيانات التدريب أو الضبط أو الاختبار المحجوز، وضوابط تسرب البيانات، وفرط التخصيص، والمعايرة، والمتانة، والانحراف.
 - اشرح المراجعة البشرية وإمكانية الاعتراض وتحديث النموذج وسجلات التدقيق والاستجابة للحوادث.
 - اذكر المعالجين الفرعيين والاستضافة ونقل البيانات وشهادات الأمن واختبارات الاختراق.
- الجوانب التجارية والتعاقدية
- عرّف مسؤوليات العميل والمورد بالاستناد إلى ISO 10667 أو إطار مكافئ عند الاقتضاء.
 - ضمّن حقوق الوصول إلى الأدلة والتدقيق، واتفاقيات مستوى الخدمة، وواجبات الإبلاغ عن الحوادث، وضبط التغيير، وإرجاع البيانات أو حذفها، وخطة الخروج.
 - حدد ما الذي يحتفظ به العميل بما يكفي لإعادة إنتاج الدرجات والدفاع عن القرارات.

ملخص قرار الاختيار

حقل	الاستنتاج
الاستخدام المقصود ومستوى المخاطر أهم نقاط القوة في الأدلة الفجوات أو القيود الجوهرية متطلبات الاجتياز المستوفاة / غير المستوفاة شروط الإطلاق متطلبات الدراسة المحلية متطلبات المراقبة	
التوصية	اعتماد / اعتماد مشروط / تجربة فقط / إعادة تصميم / رفض
مالك القرار وتاريخ المراجعة الملحق العملي C قالب ميثاق مشروع التقييم	

1. هوية المشروع

الحقل	يُستكمل للمشروع
اسم المشروع والراعي	
مالك القرار المسؤول	
مسؤول علم التقييم	
مدير المشروع	
الدول واللغات ووحدات الأعمال	
تاريخ الإطلاق المستهدف وتواريخ المراجعة	
مستوى المخاطر ومبرراته	
2. القرار والغرض	

السؤال	إجابة المشروع
ما قرار التوظيف أو التطوير الذي سيدعمه التقييم؟	
هل المخرج استشاري، أم أحد المدخلات، أم شرط اجتياز، أم ترتيب، أم محدد بدرجة كبيرة للقرار؟	
ما مشكلة القرار أو نوع الخطأ الحالي الذي يعالجه المشروع؟	
ما نتائج الأعمال ونتائج المشاركين التي تعرّف النجاح؟	
ما الاستخدامات المستبعدة صراحةً من نطاق المشروع؟	
3. السكان والسياق	
الأدوار والمستويات:	
فئة المتقدمين أو الموظفين أو المشاركين:	
قنوات التوظيف أو عمليات المواهب:	
الدول واللغات:	
الحجم المتوقع والموسمية:	
التكنولوجيا وبيئة التطبيق:	
اعتبارات الإعاقة أو اللغة أو الوصول الرقمي ذات الصلة:	
المجموعات أو التقاطعات التي تتطلب مراقبة:	
4. أدلة العمل والبنية	

المتطلب	التعريف والدليل	الأهمية / المستوى المطلوب عند الدخول	الأسلوب المقترح
---------	-----------------	--------------------------------------	-----------------

5. هندسة التقييم

المرحلة	الأسلوب	الزمن	الدرجة أو المخرج	دوره في القرار	المستخدم المسؤول

سجل نموذج الدمج، والأوزان، والعتبات، ونطاقات المراجعة، وقاعدة البيانات المفقودة، وقاعدة إعادة الاختبار، ومعالجة الحوادث، والتجاوزات المسموح بها.

6. خطة الأدلة والتجربة

سؤال الدليل	الأسلوب	العينة / المصدر	معياري النجاح	المسؤول
صلة المحتوى والعمل				
عمليات الاستجابة وسهولة الاستخدام				
الثبات والدقة				
البنية الداخلية				
العلاقة بالمعيار والقيمة الإضافية				
العدالة ونتائج المجموعات الفرعية				
إمكانية الوصول والتيسيرات				
تجربة المرشح والعمليات				
الأمن والخصوصية وتدفق البيانات				

7. رحلة المشارك

أرسم رحلة كاملة تشمل: الدعوة، والاستعداد، والفحص التقني، وطلب التيسيرات، وتسجيل الدخول، والإدارة، والدعم، والحدث، والإكمال، والنتيجة، والتغذية الراجعة، والمراجعة أو الاعتراض، والاحتفاظ، والحذف.

8. الحوكمة ومصفوفة RACI

القرار	المسؤول عن التنفيذ	المسؤول النهائي	المستشار	من يتم إبلاغه
اعتماد البنية والمخطط				
اعتماد الموزد أو الأداة				
اعتماد التسجيل والعتبات				
اعتماد الخصوصية وإمكانية الوصول				
اعتماد التجربة				
اعتماد الإطلاق التشغيلي				
اعتماد تغيير جوهري				
تعليق الأداة أو سحبها				

9. بوابة الإطلاق

- تم اعتماد الغرض والنطاق وتحليل العمل والبني.
- تمت مراجعة الأدلة الفنية والمراجعة المستقلة المناسبة.
- تمت مطابقة مواصفات التسجيل والتقارير مع حالات اختبار مستقلة.
- اختُبرت إمكانية الوصول والتيسيرات من البداية إلى النهاية.
- تم توثيق معلومات المشرح والدعم والحوادث وإعادة الاختبار ومسارات المراجعة.
- تم التحقق من الخصوصية والأمن وتدفق البيانات والاحتفاظ والتزامات المورد.
- أكمل المستخدمون التدريب وفحوص الكفاءة المطلوبة.
- تحققت معايير التجربة، أو تم توثيق الاستثناءات ومالك المخاطر.
- توجد لوحة مراقبة بمسؤولين محددين وعتبات وإجراءات وتصعيد ومؤشرات للتعليق.

10. سجل المخاطر والافتراضات والقرارات

البند	الدليل / الأساس	الخطر	الإجراء المطلوب	المسؤول	التاريخ

الملحق العملي D

مخطط تقرير التحقق / التقرير الفني

الملخص التنفيذي

- تفسيرات الدرجات والاستخدامات المقصودة.
- السكان والسياق وإصدار الأداة والمعيار وقاعدة القرار.
- النتائج الرئيسية للأدلة وعدم اليقين والقيود.
- العدالة وإمكانية الوصول والتشغيل وحماية البيانات.
- التوصية: اعتماد / شروط / تجربة فقط / إعادة تصميم / تعليق / سحب.

1. الغرض وحجة التحقق

القرار؛ والبنية؛ والسكان؛ والادعاءات؛ ومصفوفة الادعاء - الدليل - التهديد.

2. تحليل العمل والبنية

المصادر والمشاركين والمهام والحوادث الحرجة والمتطلبات والمستوى وقابلية التدريب وتعريفات البني والمؤشرات والمخطط ومؤهلات الخبراء وكيفية معالجة الخلاف.

3. الأداة والإدارة

المحتوى والنماذج والمدة وطريقة التقديم والتعليمات والتدريب والأجهزة والإشراف والتيسيرات واللغات والأمن والحوادث والانحرافات والإصدارات.

4. التسجيل والتقارير

تسجيل المفردات أو التمارين، وتدريب المقيمين، والبيانات المفقودة، ومؤشرات الجودة، والتحويلات، والمعايير، والفئات، وفترات الثقة، والدرجات المركبة، والعتبات، وقواعد التقارير، وضمان الجودة.

5. العينة وجودة البيانات

أحجام العينات ومصادرها؛ الإدراج والاستبعاد؛ التمثيل؛ المفقودات؛ الاستجابات غير المعتادة؛ تقييد المدى؛ جودة المعيار؛ ربط البيانات؛ التقسيم بين التدريب والتحقق والاختبار عند الاقتضاء.

6. التحليلات والفرضيات

ما هو تأكيد ما هو استكشافي؛ خطة التحليل؛ البرمجيات؛ الافتراضات؛ عدم اليقين؛ تعدد الاختبارات؛ القوة أو الدقة؛ والتحقق المتقاطع.

7. الثبات والدقة

لكل درجة: الاتساق الداخلي، وإعادة الاختبار، والنماذج البديلة، والثبات بين المقيمين، وقابلية التعميم، واتساق التصنيف، والدقة الشرطية، والخطأ المعياري، وفترات الدرجات.

8. البنية الداخلية وعمليات الاستجابة

الأبعاد؛ والتحليل العاملي أو نظرية الاستجابة للمفردة؛ والاعتماد المحلي؛ وسلوك المفردات؛ والمقابلات المعرفية؛ وزمن الاستجابة؛ وإدراك المقيمين؛ وسهولة الاستخدام؛ والأنماط غير المعتادة.

9. العلاقات بالمعايير والمتغيرات الأخرى

العلاقات المرصودة والمصححة؛ وجودة المعيار؛ والتصميم التنبؤي أو المتزامن؛ والقيمة الإضافية؛ والمعايرة؛ والعلاقات حسب المجموعات الفرعية؛ وعدم اليقين؛ والتحقق المتقاطع. لا تُقدّم ادعاءات سببية ما لم يسمح التصميم بذلك.

10. العدالة وإمكانية الوصول

مراجعة المحتوى؛ وإمكانية الوصول؛ والتيسيرات؛ والأداء التفاضلي للمفردة (DIF)؛ وتكافؤ القياس؛ والدقة والتنبؤ حسب المجموعات؛ ونتائج الاختيار؛ وتجربة المرشح؛ والتحليلات التقاطعية بقدر ما تسمح البيانات؛ والبدائل الأقل عبثاً أو أثراً سلبياً.

11. قواعد القرار والعواقب

العتبات؛ ونطاقات المراجعة؛ والإيجابيات والسلبيات الكاذبة؛ ومعدلات الاختيار؛ والمنفعة؛ والعبء؛ والتجاوزات؛ وسلوك المستخدم؛ والتلاعب أو التدريب المفرط.

12. الأمن والخصوصية والتكنولوجيا

تدفق البيانات؛ والوصول؛ والاحتفاظ؛ والحذف؛ والمنصة؛ وخصائص النموذج؛ وإصداراته؛ والانحراف؛ والتعرض الأمني؛ والذكاء الاصطناعي؛ والإشراف البشري؛ والحوادث.

13. الاستنتاجات والشروط والقيود

ما هو مدعوم، وما هو غير مدعوم، وما هو غير معلوم؛ وشروط الاستخدام؛ ومقاييس المراقبة وعتباتها؛ ومحفزات التعليق؛ والمسؤول؛ وتاريخ المراجعة؛ وسلطة الاعتماد.

14. الملاحق الفنية

أدلة تحليل الوظيفة؛ والمخطط؛ ومواصفات المفردات؛ وأدوات الدراسة؛ والبرمجيات أو صياغة التحليل؛ وقاموس البيانات؛ والجداول؛ وسجل القرارات؛ وتعليقات المراجعين؛ والاعتمادات؛ وسجل الإصدارات؛ مع مراعاة أمن الاختبار والسرية.

الملحق العملي E

قالب التواصل مع المرشح

النص التالي نقطة انطلاق. عدّله بما يتفق مع العملية الفعلية، والقانون، وإشعار الخصوصية، ومسار إمكانية الوصول، ونبرة المؤسسة. لا تعد بتغذية راجعة أو سرية أو مراجعة بشرية لا تستطيع العملية فعلياً توفيرها.

الموضوع: معلومات حول تقييمك

شكراً لمشاركتك في التقييم الخاص بـ[الدور أو البرنامج]. صُمم التقييم لجمع أدلة موحدة حول [البنى أو الأنشطة بصياغة مبسطة] ذات الصلة بـ[متطلبات الدور المحددة أو غرض التطوير].

كيف ستستخدم النتيجة

ستستخدم نتيجتك كأحد المدخلات / كشرط للفرز / لدعم التطوير / غير ذلك. وسيتم النظر فيها إلى جانب [أساليب أخرى]. سيقوم / لن يقوم التقييم باتخاذ قرار آلياً. [اشرح دور المراجعة البشرية وكيف يمكن الإبلاغ عن الأخطاء الجوهرية.]
ما الذي تتوقعه

- أنشطة التقييم: [وصف مختصر من دون كشف محتوى آمن أو مقيد].
- الزمن التقريبي للإكمال: [الوقت]، بما في ذلك [التدريب أو فترات الراحة].
- نافذة الإكمال والموعود النهائي: [التواريخ والمنطقة الزمنية].
- المواد أو الأدوات المسموح بها: [قائمة].
- متطلبات الجهاز أو المتصفح أو الصوت أو الكاميرا أو الاتصال: [قائمة أو رابط فحص].
- الإشراف أو التحقق من الهوية: [اشرح بوضوح وبصورة متناسبة].

الاستعداد

لا تحتاج إلى تدريب متخصص. نوصي بأن تراجع مواد التدريب، وتختار مكاناً هادئاً، وتفحص الجهاز، وتخصص وقتاً كافياً. تتوفر مواد التدريب في [المكان]. توضح مواد التدريب صيغة التقييم ولا تدخل في التسجيل.

إمكانية الوصول والتيسيرات

نرغب في توفير فرصة مكافئة لإظهار المتطلبات ذات الصلة. لطلب تيسير أو مناقشة حاجة تتعلق بإمكانية الوصول، تواصل مع [مسار سري] بحلول [تاريخ إن أمكن]. لا تحتاج إلى الإفصاح عن معلومات تتجاوز ما هو ضروري لتحديد تيسير فعال. ولن يؤثر طلب التيسير سلباً في طلبك.

البيانات والخصوصية

سنجمع [فئات البيانات]. نستخدم البيانات لأغراض [الأغراض]، ونشارك مع [المستلمين]، ونخزن في [المواقع] إذا كان الإفصاح مطلوباً، ويحتفظ بها لمدة [الفترة أو القاعدة]. [اشرح الأتمتة والتسجيل والتحقق ومراقبة المساواة والاستخدام البحثي ذات الصلة.] تتوفر معلومات الخصوصية الكاملة في [المكان].

حادث تقني أو ظرف آخر

إذا كان انقطاع أو مشكلة في الجهاز أو حالة صحية أو فشل في التيسير أو أي حدث آخر قد أثر في الأداء، فتوقف متى كان ذلك ممكناً وتواصل مع [مسار الدعم]. سنراجع الحادث وفق عملية المعالجة وإعادة الاختبار الموثقة. لا تنشئ حساباً ثانياً ولا تعد الاختبار ما لم يُطلب منك ذلك.

النتيجة والتغذية الراجعة والمراجعة

سيتم إبلاغك بـ[الخطوة التالية / النتيجة] بحلول [الوقت]. [صف التغذية الراجعة]. إذا كنت تعتقد أن معلومات جوهرية كانت غير صحيحة أو أن العملية لم تُدار كما وُصفت، فتواصل مع [مسار المراجعة] خلال [الفترة]. لا يمكن كشف محتوى المفردات ومفاتيح التسجيل إذا كان ذلك يهدد أمن الاختبار، لكن يمكن مراجعة العملية وكيفية استخدام النتيجة.

بيانات التواصل

دعم التقييم: [جهة الاتصال] إمكانية الوصول: [جهة الاتصال] الخصوصية: [جهة الاتصال] فريق التوظيف أو البرنامج: [جهة الاتصال]
الملحق العملي F

قائمة فحص تدقيق جودة التقييم

قيم كل بند: نعم، جزئياً، لا، غير منطبق. سجّل الدليل، والمسؤول، والإجراء. تعني «نعم» أن الدليل الحالي تمت مراجعته فعلياً، وليس أن شخصاً يعتقد أن الضابط موجود.

الغرض والاستخدام

- القرار والغرض والسكان والعواقب موثقة.
- الاستخدامات المدعومة والمحظورة صريحة.
- يعالج التقييم فجوة أدلة محددة بدلاً من تكرار الأساليب الحالية.
- تمت الموافقة على دور التقييم في القرار، والأوزان، والعتبات، ونطاقات المراجعة، والتجاوزات.
- مستوى المخاطر حديث ويعكس الأثمة والحجم والحساسية وإمكانية الاعتراض.

صلة الوظيفة والبنية

- تدعم أدلة تحليل العمل الحالية المتطلبات التي يتم تقييمها.
- البنى معرفة بما يشمل ما يدخل فيها وما يُستبعد منها والمؤشرات السلوكية.
- يربط المخطط المحتوى والأساليب بالمتطلبات والمستوى.
- تمت مراجعة نقص البنية والتباين غير ذي الصلة بالبنية.
- يمثل نموذج المعيار جوانب الأداء المهمة والعوامل السياقية.

أدلة الأداة

- يتوفر الدليل الفني الحالي والمراجعات المستقلة.
- النبات والدقة مناسبان لكل درجة يتم الإبلاغ عنها.
- تدعم أدلة الصدق التفسير والسكان والاستخدام الفعلي.
- المجموعات المعيارية أو المستويات ذات صلة وحديثة وموثقة.
- حدود القرار والفئات لها مبررات وتحليل للأخطاء.
- النسخ المكيفة أو المترجمة لديها أدلة على التكافؤ.

العدالة وإمكانية الوصول

- خضع المحتوى لمراجعة التحيز واللغة والثقافة وإمكانية الوصول.
- اختُبر التقديم الرقمي وفق ممارسات إمكانية الوصول الحديثة.
- يوجد مسار سري لطلب التيسير وعملية قرار موثقة.
- تتم مراقبة الإكمال والدرجات والاختيار والمعايير حسب المجموعات الفرعية.
- تُستخدم تحليلات DIF أو تكافؤ القياس أو التنبؤ عندما تكون مناسبة وممكنة.
- جرى النظر في بدائل أقل أثراً سلبياً أو أقل عبئاً.
- توجد مسارات ذات معنى للمراجعة والتصحيح والمعالجة للمشاركة.

الإدارة والتجربة

- التعليمات والتدريب والوقت والبيئة والأدوات والدعم موحدة.
- الإشراف وضوابط الهوية متناسبة وشفافة.
- قواعد الحوادث والانقطاع والإبطال وإعادة الاختبار والاعتراض موثقة.

- تشرح اتصالات المرشح بدقة الغرض والاستخدام والبيانات والخطوات التالية.
- تتم مراقبة التجربة والانسحاب والدعم والشكاوى حسب المجموعات ذات الصلة.

التسجيل والتقارير

- تحدد مواصفة موقعة جميع المفاتيح والأوزان والتحويلات وقواعد البيانات المفقودة والمعايير والتصنيفات.
- تتحقق اختبارات الوحدة والمطابقة المستقلة من صحة التسجيل.
- تتم مراقبة تدريب المقيمين ومعايرتهم ودرجة الاتفاق بينهم.
- تحدد التقارير الغرض والإصدار والمعيار والدقة والقيود والجمهور المقصود.
- يظل التفسير السردى منضبطاً ويتجنب الادعاءات الحتمية أو السريية.
- يستند الوصول إلى التقارير الفردية إلى الكفاءة والحاجة المهنية.

البيانات والخصوصية والأمن والذكاء الاصطناعي

- رُسم تدفق البيانات كاملاً من الجمع حتى الحذف.
- الجمع ضروري ومتناسب ومدعوم بغرض واضح.
- ضوابط الوصول والتشفير والتسجيل والاحتفاظ والحذف والمعالجين الفرعيين فعالة.
- المحتوى المقيد والمفاتيح والنماذج وبنوك المفردات محمية وتتم مراقبة تعرضها.
- خصائص الذكاء الاصطناعي وبيانات التدريب والتسميات والمتغيرات الجوهرية موثقة.
- يتم تقييم تسرب البيانات وفطرت التخصيص والأداء حسب المجموعات والمعايير والمتانة والانحراف.
- يستطيع المراجعون البشريون فهم المخرجات واستجوابها وتصحيحها وتجاوزها بصورة مناسبة.
- توجد خطط استجابة مجربة لحوادث الأمن والخصوصية ونزاهة الاختبار والنماذج.

الحوكمة ودورة الحياة

- الأداة مسجلة في مخزون التقييمات مع المالك والحالة.
- يحمل المستخدمون المؤهلات الحالية وأكملوا التدريب والمعايرة المطلوبة.
- تدعم العقود التزامات المورد والوصول إلى الأدلة والتدقيق والإخطار بالتغيير والخروج.
- تحمل كل درجة وتقرير تشغيليين معرف إصدار.
- تؤدي التغييرات الجوهرية إلى مراجعة أثر وإعادة تحقق متناسبة.
- للمراقبة مسؤولون محددون وعتبات وإجراءات تصحيح ومحفزات تعليق.
- تتم مراجعة التجاوزات والاستثناءات والحوادث والشكاوى والإجراءات العلاجية.
- توجد خطة لسحب الأداة والخروج من البيانات.
- تم تسجيل تاريخ المراجعة الرسمية التالية وصاحب الاعتماد المسؤول.

سجل نتائج التدقيق

النتيجة	الدليل	الخطر	الإجراء المطلوب	المسؤول	تاريخ الاستحقاق	التحقق
---------	--------	-------	-----------------	---------	-----------------	--------

الملحق العملي G

المسرد

المصطلح	التعريف العملي
التيسير Accommodation	تعديل يهدف إلى إزالة حاجز مرتبط بالإعاقة أو حاجز مشروع آخر مع الحفاظ قدر الإمكان على البنية المستهدفة.
الأثر السلبي Adverse impact	اختلاف جوهري في نتيجة الاختيار لمجموعة، يُقيم وفق الإطار القانوني والسياساتي ذي الصلة.
التقييم الخوارزمي Algorithmic assessment	إجراء تحول فيه قواعد أو نماذج حاسوبية البيانات إلى درجات أو توصيات أو ترتيبات أو قرارات.
ثبات النماذج البديلة Alternate-form reliability	اتساق الدرجات عبر نماذج يُقصد أن تكون متكافئة.
انطباع المتقدم Applicant reaction	تصور المشارك أو تجربته بشأن الصلة والعدالة والشفافية والعبء وطريقة المعاملة.
التقييم Assessment	العملية الأوسع لجمع الأدلة ودمجها وتفسيرها واستخدامها، وقد تشمل اختبارات ومقابلات ومحاكاة وحكماً مهنيّاً.
مركز التقييم Assessment centre	عملية موحدة متعددة الأساليب تتضمن تمارين متعددة ومقيمين مدربين، وليست مجرد مكان مادي.
معدل الأساس Base rate	معدل انتشار النتيجة المستهدفة أو النجاح في السكان المعنيين قبل استخدام التقييم.
مؤشر سلوكي Behavioral anchor	وصف ملموس لدليل سلوكي قابل للملاحظة يرتبط بمستوى تقييم معين.
التحيز Bias	تشويه منهجي في القياس أو التنبؤ أو العملية. يجب تحديد نوع التحيز والتحقيق فيه.
المخطط Blueprint	خطة تربط مجالات البنية أو المحتوى بالمفردات والتمارين والصيغ والصعوبة والتسجيل والتغطية.
المعايرة Calibration	مواءمة مقاييس الدرجات أو معاملات المفردات أو النماذج أو المقيمين؛ وفي التنبؤ، مدى تطابق الاحتمالات المتنبأ بها مع النتائج المرصودة.
المرشح Candidate	شخص يجري النظر فيه لقرار توظيف أو تطوير أو اعتماد أو موهبة.
النظرية الكلاسيكية للاختبار Classical test theory	إطار قياس تُهم فيه الدرجة المرصودة باعتبارها مكوناً من الدرجة الحقيقية مضافاً إليه الخطأ.
اتساق التصنيف Classification consistency	احتمال حصول الشخص على الفئة أو قرار الاجتياز/عدم الاجتياز نفسه عبر تكرارات متكافئة.
الدرجة المركبة Composite score	درجة تنشأ من دمج مكونين أو أكثر موحدتين أو متوافقين وفق أوزان محددة.
الاختبار التكيفي المحوسب Computerized adaptive testing	اختبار رقمي تُختار فيه المفردة التالية جزئياً بناءً على الاستجابات السابقة لاستهداف المعلومات بكفاءة.
فترة الثقة Confidence interval	نطاق يعبر عن عدم اليقين حول تقدير وفق أسلوب إحصائي محدد.
أدلة الصدق المرتبطة بالعواقب Consequential validity evidence	أدلة حول الآثار المقصودة وغير المقصودة المرتبطة باستخدام الدرجات وقواعد القرار.
البنية Construct	خاصية نظرية يُستدل عليها من الملاحظات، مثل الاستدلال أو الضمير المهني أو الحكم.
نقص البنية Construct deficiency	عدم أخذ عينة كافية من أجزاء مهمة في البنية أو مجال المعيار المقصود.
التباين غير ذي الصلة بالبنية Construct-irrelevant variance	تباين في الدرجات تسببه عوامل خارج البنية المراد قياسها.
أدلة صدق المحتوى Content validity evidence	أدلة على أن محتوى التقييم يمثل بصورة كافية مجال الوظيفة أو البنية المحدد.
المعيار Criterion	نتيجة أو مؤشر يُستخدم لتقييم التقييم، مثل الأداء الوظيفي أو التعلم أو السلامة أو التقدم.

المصطلح	التعريف العملي
تلوث المعيار Criterion contamination	تأثر درجات النتيجة بعوامل غير ذات صلة أو بمعرفة مسبقة بالمتنبئ.
نقص المعيار Criterion deficiency	إغفال المعيار لجوانب مهمة من النتيجة محل الاهتمام.
أدلة الصدق المرتبطة بالمعيار Criterion-related validity	أدلة حول العلاقة بين درجات التقييم والنتائج ذات الصلة.
التحقق المتقاطع Cross-validation	تقييم نموذج أو قاعدة قرار على بيانات لم تُستخدم في تطويره أو ضبطه.
حد القرار Cut score	عتبة تُستخدم للتصنيف أو الاجتياز أو عدم الاجتياز أو توجيه المشاركين لمسارات مختلفة.
تسرب البيانات Data leakage	دخول معلومات غير متاحة عند نقطة القرار أو مشتقة بصورة غير صحيحة من النتيجة إلى تطوير النموذج أو التسجيل.
قاعدة القرار Decision rule	منطق محدد مسبقاً لدمج الأدلة في توصية أو تصنيف أو إجراء.
ال أداء التفاضلي للمفردة Differential item functioning (DIF)	سلوك المفردة بصورة مختلفة بين مجموعات متساوية على السمة الكامنة بحسب النموذج. هو إشارة للتحقيق وليس دليلاً آلياً على التحيز.
الانحراف Drift	تغير في السكان أو البيانات أو تعرض المفردات أو النموذج أو المعيار أو العلاقات قد يضعف معنى الدرجة بعد التطبيق.
حجم الأثر Effect size	وصف كمي لحجم الفرق أو العلاقة، وهو مختلف عن الدلالة الإحصائية.
المعادلة أو الربط Equating or linking	أساليب إحصائية لوضع درجات النماذج أو الإصدارات المختلفة على مقياس قابل للمقارنة.
نموذج الأدلة Evidence model	وصف للاستجابات المرصودة التي من شأنها دعم استنتاج بشأن بنية معينة.
الصدق الظاهري Face validity	مدى ظهور التقييم للمستخدمين على أنه ذو صلة أو معقول؛ يفيد تجربة المستخدم لكنه ليس دليلاً على الصدق الفني.
التحليل العاملي Factor analysis	أساليب إحصائية لفحص أو اختبار أنماط التباين المشترك بين المفردات أو الدرجات.
العدالة Fairness	تقييم واسع للوصول والمعاملة والقياس والتنبؤ والقرارات والعواقب والشفافية والقيم.
سليبي كاذب False negative	شخص كان سيحقق المعيار لكنه رُفض أو صُنّف سلبياً.
إيجابي كاذب False positive	شخص لا يحقق المعيار لكنه اختير أو صُنّف إيجابياً.
مفردة الاختيار الإجباري Forced-choice item	صيغة تتطلب الاختيار أو الترتيب بين بدائل، وتُستخدم كثيراً في تقييم الشخصية.
قابلية التعميم Generalizability	مدى بقاء الأدلة أو الدرجات قابلة للتطبيق عبر الأشخاص أو الوظائف أو المواقف أو المقيمين أو المفردات أو الزمن ذات الصلة.
تقييم عالي المخاطر High-stakes assessment	تقييم يمكن أن تؤثر نتيجته مادياً في التوظيف أو الدخل أو التقدم أو الاعتماد أو نتيجة أخرى ذات عواقب مهمة.
الإشراف البشري Human oversight	مراجعة كفؤة ومستنيرة وممكنة وقابلة للتدقيق تستطيع اكتشاف أخطاء النظام الجوهرية وتصحيحها.
الصدق الإضافي Incremental validity	علاقة مفيدة إضافية بمعيار بعد أخذ الأساليب أو المعلومات الحالية في الاعتبار.
الاتساق الداخلي Internal consistency	درجة ترابط المفردات التي تسهم في درجة معينة وفق نموذج محدد.
الثبات بين المقيمين Inter-rater reliability	اتساق أو اتفاق المقيمين الذين يقيمون الدليل نفسه.
درجة إبساطية Ipsative score	درجة تُعرّف بالنسبة إلى درجات الشخص الأخرى، ما يخلق غالباً اعتماداً يحد من المقارنة بين الأشخاص.
بنك المفردات Item bank	مجموعة مضبوطة من مفردات معايرة أو مضمونة الجودة تُبنى منها النماذج أو الاختبارات التكيفية.

المصطلح	التعريف العملي
تمييز المفردة Item discrimination	مدى قدرة المفردة على التمييز بين الأشخاص عند مستويات مختلفة من السمة التي يمثلها النموذج.
نظرية الاستجابة للمفردة Item response theory (IRT)	نماذج تربط احتمال الاستجابة للمفردة بسمة كامنة وخصائص المفردة.
تحليل الوظيفة Job analysis	دراسة منهجية للمهام والسياق والمتطلبات وأدلة الأداء تُستخدم لتعريف محتوى التقييم والقرارات.
التحقق المحلي Local validation	أدلة يتم توليدها داخل مؤسسة المستخدم أو مكانه أو وظيفته أو تطبيقه الفعلي.
خطأ القياس Measurement error	تباين في الدرجة لا يرتبط بالفروق المستقرة المقصودة وفق الشروط المحددة.
تكافؤ القياس Measurement invariance	دليل على أن نموذج القياس يعمل بدرجة متقاربة كفاية عبر مجموعات أو مناسبات مختلفة بحيث تكون المقارنات المقصودة مبررة.
معايرة النموذج Model calibration	مدى اتفاق احتمالات النتائج المتنبأ بها مع النتائج المرصودة، أو دقة تفسير الدرجات.
المجموعة المعيارية Norm group	العينة المرجعية المستخدمة لتفسير درجة مقارنة الآخرين.
درجة مرجعية معيارياً Norm-referenced score	درجة تُفسر بالمقارنة مع مجموعة مرجعية.
الرتبة المئينية Percentile rank	نسبة أفراد المجموعة المرجعية الذين حصلوا على درجة مساوية أو أقل؛ المئينات ليست وحدات متساوية الفواصل.
التحيز التنبؤي Predictive bias	فروق منهجية في علاقات التنبؤ تؤدي إلى زيادة أو نقص التنبؤ لمجموعات بعينها.
دراسة الصدق التنبؤي Predictive validity study	تصميم تُجمع فيه بيانات التقييم قبل معيار لاحق، ويفضل ألا تؤثر درجات التقييم في المعيار.
السيكومتريا Psychometrics	علم وممارسة القياس النفسي والتربوي.
الثبات Reliability	اتساق أو دقة الدرجات في شروط محددة.
عملية الاستجابة Response process	العملية المعرفية أو السلوكية أو الحكمية المستخدمة لإنتاج استجابة لمفردة أو تقييم.
نمط الاستجابة Response style	طريقة منهجية في استخدام فئات الاستجابة قد تكون مستقلة جزئياً عن محتوى المفردات.
نطاق المراجعة Review zone	نطاق درجات حول عتبة تُستخدم فيه أدلة منظمة إضافية أو مراجعة إضافية.
درجة محولة Scaled score	درجة محولة إلى مقياس تقرير محدد.
إجراء الاختيار Selection procedure	أي أسلوب يُستخدم لاتخاذ أو دعم قرار توظيف، بما في ذلك الاختبارات والمقابلات ومتطلبات الخبرة والخوارزميات والمحاكاة.
نسبة الاختيار Selection ratio	نسبة المتقدمين أو المرشحين الذين تم اختيارهم.
الحساسية Sensitivity	نسبة الإيجابيات الفعلية التي تم تحديدها بصورة صحيحة.
النوعية Specificity	نسبة السلبيات الفعلية التي تم تحديدها بصورة صحيحة.
الخطأ المعياري للقياس Standard error of measurement	تقدير لعدم اليقين في الدرجة معبر عنه بوحدات مقياس الدرجة المبلغ عنها.
التوحيد Standardization	ضبط المحتوى والتعليمات والظروف والتسجيل والتفسير ذات الصلة لدعم قابلية المقارنة.
المقابلة المنظمة Structured interview	مقابلة مرتبطة بالوظيفة تستخدم أسئلة واستقصاءات ومؤشرات تسجيل متسقة، ومحاورين مدربين، وقواعد دمج واضحة.
خبير الموضوع Subject-matter expert	شخص لديه معرفة حديثة وذات صلة بالعمل أو المجال أو السكان ويسهم بأدلة منظمة.
الاختبار Test	أداة أو إجراء موحد يستثير عينة من الاستجابات أو السلوك ويقوم بتسجيلها.

المصطلح	التعريف العملي
ثبات إعادة الاختبار Test-retest reliability	استقرار الدرجات عبر فترة زمنية مناسبة عندما لا يُتوقع تغير حقيقي.
درجة T T-score	درجة معيارية تُعرّف عادةً بالصيغة $10z + 50$.
المنفعة Utility	القيمة العملية للتقييم بالنظر إلى النتائج والأخطاء والتكلفة والحجم ومدة البقاء والعدالة والتنفيذ.
الصدق Validity	مدى دعم الأدلة والنظرية لتفسيرات الدرجات واستخداماتها المقصودة.
التحقق Validation	العملية المخططة لبناء الأدلة والحجة الداعمة للاستخدام وتقييمها وتحديثها.
عينة العمل Work sample	مهمة موحدة تشبه نشاطاً وظيفياً مهماً إلى حد كبير وتُسجل وفق قواعد محددة.
درجة z z-score	موقع معياري يُعبّر عنه بوحدات الانحراف المعياري عن متوسط مرجعي.

الملحق العملي H

معايير وإرشادات وقراءات مختارة

المعايير الأساسية للاختبارات

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). Standards for Educational and Psychological Testing. <https://www.testingstandards.net/open-access-files.html>

Society for Industrial and Organizational Psychology. (2018). Principles for the Validation and Use of Personnel Selection Procedures (5th ed.). <https://www.apa.org/ed/accreditation/personnel-selection-procedures.pdf>

International Organization for Standardization. (2020). ISO 10667-1:2020 Assessment service delivery - Procedures and methods to assess people in work and organizational settings - Part 1: Requirements for the client. <https://www.iso.org/standard/74716.html>

International Organization for Standardization. (2020). ISO 10667-2:2020 Assessment service delivery - Procedures and methods to assess people in work and organizational settings - Part 2: Requirements for service providers. <https://www.iso.org/standard/74717.html>

إرشادات اللجنة الدولية للاختبارات (ITC)

International Test Commission. (2001). The ITC Guidelines on Test Use. <https://www.intestcom.org/page/15>

International Test Commission. (2017). The ITC Guidelines for Translating and Adapting Tests (2nd ed.). https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf

International Test Commission. (2014). Guidelines on Quality Control in Scoring, Test Analysis, and Reporting of Test Scores. <https://www.intestcom.org/page/17>

International Test Commission. (n.d.). The ITC Guidelines on the Security of Tests, Examinations, and Other Assessments. <https://www.intestcom.org/page/18>

International Test Commission and Association of Test Publishers. (2022). Guidelines for Technology-Based Assessment. <https://www.intestcom.org/page/16>

إرشادات اختيار الأفراد والذكاء الاصطناعي

Society for Industrial and Organizational Psychology. (2023). Considerations and Recommendations for the Validation and Use of AI-Based Assessments for Employee Selection. <https://www.siop.org/wp-content/uploads/2024/06/Considerations-and-Recommendations-for-the-Validation-and-Use-of-AI-Based-Assessments-for-Employee-Selection-January-2023.pdf>

National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). <https://www.nist.gov/itl/ai-risk-management-framework>

European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>

European Commission. (2026). Guidelines for providers and deployers of AI high-risk systems. <https://digital-strategy.ec.europa.eu/en/policies/guidelines-ai-high-risk-systems>

Information Commissioner's Office. (2024). AI tools used in recruitment: outcomes report. <https://ico.org.uk/action-weve-taken/audits-and-overview-reports/2024/11/ai-tools-used-in-recruitment/>

Information Commissioner's Office. (2026). Recruitment rewired: an update on the ICO's work on recruitment and automation. <https://ico.org.uk/about-the-ico/what-we-do/recruitment-rewired/>

U.S. Equal Employment Opportunity Commission. (2007). Employment Tests and Selection Procedures. <https://www.eeoc.gov/laws/guidance/employment-tests-and-selection-procedures>

World Wide Web Consortium. (2024). Web Content Accessibility Guidelines (WCAG) 2.2. <https://www.w3.org/TR/WCAG22/>

British Psychological Society, Psychological Testing Centre. (n.d.). Qualifications in test use and the Register of Qualifications in Test Use. <https://www.psychtesting.org.uk/>

تحديث الأدلة والمنظور البحثي

Sackett, P. R., Zhang, C., Berry, C. M., and Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*, 107(11), 2040-2068. <https://doi.org/10.1037/apl0000994>

الغرض من إدراج هذه المقالة منهجي، وليس استبدال ترتيب ثابت بترتيب ثابت آخر. تعتمد تقديرات التحليل التجميعي على العينات والوظائف والمعايير وجودة الدراسات والتصحيحات والافتراضات. ينبغي للممارسين استخدام البحث المتراكم لتكوين فرضيات مستنيرة، ثم تقييم السكان الفعليين وطريقة التطبيق والقرار.

تسلسل تعلم مقترح

1. ابدأ بـ *Standards for Educational and Psychological Testing* وإرشادات ITC لاستخدام الاختبارات.
2. ادرس مبادئ SIOP للتحقق المرتبط بالعمل ومعيار ISO 10667 لمسؤوليات العميل ومقدم الخدمة.
3. أضف إرشادات ITC الخاصة بالتكنولوجيا والتكييف وضبط الجودة والأمن إلى ممارسات التنفيذ.
4. اقرأ أحدث إرشادات المساواة والإعاقة والخصوصية والذكاء الاصطناعي في كل ولاية قضائية تُستخدم فيها الأداة.
5. اقرأ الأدلة الفنية بنظرة نقدية وقارن ادعاءاتها بالبحوث المستقلة.
6. ابن كفاءة عملية تحت الإشراف من خلال تحليل الوظيفة، والأساليب المنظمة، والتفسير، والتحقق، والمراقبة، والحوكمة.

ملاحظة insyt2 - الحد المهني النهائي

هذا الدليل تعليمي. وهو ليس استشارة قانونية، ولا ترخيصاً مهنيًا، ولا مؤهلاً صادراً عن ناشر، ولا بديلاً عن الكفاءة المكتسبة تحت الإشراف. تحقق من المعايير والقانون وقواعد الأداة والأدلة المحلية الحالية قبل اتخاذ قرارات تؤثر في الناس.

المعيار المهني من insyt2

اعرف الحدود. احم القرار. واجعل الحكم المهني مرئياً.

يجب أن يخلق التقييم قرارات أفضل - لا يقيناً زائفاً.

دليل عملي لمتخصصي الموارد البشرية الذين يريدون استخدام السيكموتريا بثقة أكبر، ودليل أقوى، وحكم مهني أفضل.

يجمع هذا الإصدار المهني للشرق الأوسط بين العلم والحكمة والحرفة العملية للتقييم المهني. وقد صُمم ليكون مرجع عمل: شيئاً تستخدمه في حوارات الموردين، وتصميم المشروعات، وتعود إليه عندما ترتفع المخاطر.

استكشف حلول تقييم insyt2 عبر بوابة التقييم التابعة لـ HR Delivery

<https://insyt2-hr-assessment-centre.ynvmkxxnh6.chatgpt.site/?entry=assessment-portal#assess>

اسم المستخدم: Portal · كلمة المرور: Portal

