2025/2024

سمينار شباب الباحثين

موجز السياسات Policy Brief

"معايير تقييم جودة البيانات بالتطبيق علي مصر"

إعداد

أ. سالمة متولي عبده علي

مدرس مساعد بمركز السياسات الاقتصادية الكلية

المنسق

د. مي عوض

مدرس إدارة الأعمال

المنسق المشارك

د. إسلام خليل

مدرس الاقتصاد

معهد التخطيط القومي

إبريل 2025

إلى أي مدى يمكن أن تساهم زيادة الاعتماد على التكنولوجيا والذكاء الاصطناعي ومصادر البيانات غير التقليدية في تحسين جودة البيانات وتعزيز النمذجة في مصر؟

مقدمة

تُعد جودة البيانات عاملًا حاسمًا في نجاح جهود التنمية على مختلف المستويات. فالبيانات غير الدقيقة أو غير المحدثة تُؤدي إلى قرارات خاطئة، وهدر في الموارد، وضعف في تقييم الأثر. بينما تُمكّن البيانات عالية الجودة من التخطيط السليم، والتنفيذ الفعّال، والتقييم الموضوعي للبرامج والمشروعات التنموية.

إن التنمية الحقيقية لا تقوم على الحدس أو التقديرات، بل على أدلة واضحة تعكس الواقع بدقة. ولهذا فإن جودة البيانات – من حيث الدقة، والموثوقية، والشمول، والتوقيت – تُشكل الأساس الذي تُبنى عليه السياسات والاستراتيجيات الفعالة.

وفي سياق تحقيق أهداف التنمية المستدامة(SDGs) ، تُعد البيانات الجيدة أداة لا غنى عنها لرصد التقدم، وتحديد الفجوات، وتصميم التدخلات المناسبة. وبالتالي، فإن الاستثمار في تحسين جودة البيانات الوطنية يُعد استثمارًا مباشرًا في مستقبل التنمية.

خيارات السياسات لتحسين جودة البيانات في مصر:

- **السياسة 1: إنشاء هيئة موحدة لحوكمة البيانات, تحديث وتفعيل الاستراتيجية الوطنية للإحصاء.**
- **السياسة 2: تحديث منظومة الحسابات القومية.**
- **السياسة 3: تعزيز الشفافية وإتاحة البيانات المفتوحة.**
- **السياسة 4: الاستفادة من التكنولوجيا والابتكار في جمع وتحليل البيانات.**
- **السياسة 5: تحسين استخدام السجلات الإدارية للأغراض الإحصائية.**
- **السياسة 6: تنفيذ مبادرات بناء القدرات وبرامج المساعدة الفنية التي تقودها المنظمات الاقليمية مثل الاسكوا، صندوق النقد الدولي ،البنك الدولي.**

تعريف جودة البيانات وأبعادها المختلفة

تشكل جودة البيانات حجر الزاوية في تحديد قيمتها الحقيقية والاقتصادية ودرجة ثقة مستخدمها ولذا فإنه من الضروري أن تمتلك الهيئات الإحصائية منظومة متكاملة لإدارة وضمان الجودة تضمن إنتاج بيانات ذات قيمة عالية، موضوعية ولا تخضع للتجاذبات السياسية، ولا تخدم مصالح مجموعات الضغط، وتعد وفق المبادئ الأساسية للإحصاءات الرسمية للامم المتحدة. ولبناء نظام جودة ذي مصداقية فإن الهيئات الإحصائية عليها أولاً صياغة استراتيجية توضح فيها الغايات والقيم والرسالة والأهداف والبرامج، بالإضافة إلى مجموعة من مؤشرات الأداء التي تسمح بالتقييم الجاد للاستراتيجية الإحصائية وفق المبادئ الأساسية للعمل الإحصائي. وبالرغم من تداول تعريف الجودة وفق نظام الأيزو 9000:2015 والذي يرى أنها"درجة استيفاء مجموعة من الخصائص المتأصلة في المنتج" فإن البيانات الإحصائية تتطلب تعريفًا للجودة يتسق مع خصائصها ويركز على درجة صلاحيتها للاستخدام أو للغرض الذي تم على أساسه إنتاج هذه البيانات. وهذا التعريف يجعل الجودة تتحدد أساسًا بدرجة استيفاء احتياجات المستخدمين للبيانات، علمًا بأن حاجات المستخدمين تختلف فيما بينهم بشكل كبير. جودة الإحصاءات مسألة متعددة الجوانب ولا يوجد معيار واحد يحددها بشكل موضوعي. وعليه فإن تقييم الجودة يتم عبر تحديد مجموعة من الأبعاد التي تميزها والتي تسمح بقياسها مثل درجة الملاءمة، والدقة، والموثوقية، وحسن التوقيت، والالتزام بالمواعيد، وإمكانية وسهول الحصول عليها والوضوح، والتماسك، والقابلية للمقارنة.

الشكل رقم (1) أبعاد الجودة

تحليل الوضع الحالي لجودة البيانات في مصر:

الجدول رقم (1) : تقييم الوضع الحالي لجودة البيانات في مصر

	الوضع الحالي في مصر
الجهات المسؤولة عن إنتاج البيانات	• الجهاز المركزي للتعبئة العامة والاحصاء • الوزارات والهيئات الحكومية
الإطار المؤسسي	• قانون حماية البيانات الشخصية (رقم 151 لسنة 2020) • قانون مكافحة جرائم تقنية المعلومات (رقم 175 لسنة 2018) • قانون حرية تداول المعلومات (لم يصدر بعد بشكل نهائي) رغم النص عليه في المادة 68 من دستور 2014 • قوانين ولوائح أخرى ذات صلة • قانون العقوبات المصري: يحتوي على مواد تعاقب على انتهاك سرية البيانات أو إفشاء أسرار العمل. • قوانين تنظيم الاتصالات (قانون 10 لسنة 2003): تضمن حماية خصوصية المستخدمين ومراقبة استخدام البيانات من قبل شركات الاتصالات.
التنسيق بين الجهات	• تنسيق محدود أو غير منهجي بين الجهات المنتجة للبيانات. • تعدد مصادر البيانات
وحدات الجودة	لا توجد وحدات دائمة لإدارة الجودة داخل أغلب المؤسسات.
المعايير المستخدمة	تفاوت في المعايير والمنهجيات المستخدمة بين الجهات.
الرقمنة والتكامل	قواعد بيانات مشتتة ومنفصلة، كثير منها غير رقمي.
الثقة في البيانات	ثقة جزئية من المستخدمين بسبب التفاوت في الجودة والحداثة.

التشريعات المنظمة	غياب تشريعات واضحة وشاملة لضبط جودة البيانات ومشاركتها.
توثيق البيانات(Metadata)	التوثيق غير منتظم أو غير متاح غالبًا.
التدريب وبناء القدرات	ضعف في تأهيل الكوادر بمجال تحليل البيانات وضمان الجودة.
استخدام التكنولوجيا	استخدام محدود لتقنيات الذكاء الاصطناعي والتحليل التنبؤي.

- مشاكل جودة البيانات في مصر

الجدول رقم (2) : مشاكل تتعلق بجودة البيانات في مصر

نوع المشكلة	الوصف
بيانات غير دقيقة	الاعتماد على تقديرات أو مصادر غير موثوقة دون مسوح ميدانية .
بيانات غير مكتملة	غياب تغطية شاملة لكل مكونات الظاهرة محل الدراسة
بيانات قديمة	عدم تحديث البيانات بشكل دوري ما يؤدي إلى فقدانها الصلاحية لاتخاذ قرارات دقيقة
بيانات مكررة	تكرار البيانات لنفس الأفراد أو الكيانات بسبب عدم تكامل قواعد البيانات .
بيانات غير متناسقة	اختلاف المفاهيم والتعاريف بين الجهات المنتجة للبيانات، مما يسبب تعارضاً في النتائج
بيانات زائدة (overload)	وجود كم هائل من البيانات يصعب تحليله بكفاءة، ما يؤدي إلى إرهاق الموارد وزيادة احتمالات الخطأ(تحليل بيانات مو اقع التواصل الاجتماعي).
بيانات قديمة مكررة غير مكتملة غير متناسقة	بيانات ضعيفة الجودة

- فجوة البيانات في مصر:

تشير فجوة البيانات إلى النقص في البيانات المتاحة، أو عدم دقتها، أو عدم تحديثها، أو عدم سهولة الوصول إليها. وتصل فجوة البيانات في مصر إلي 54%.

الجدول رقم (3): نسب البيانات غير المتوفرة في قاعدة بيانات مؤشرات التنمية المستدامة

	نسبة البيانات غير المتوفرة متوسط 2000-2024 (1) تقرير التنمية المستدامة الذي ينشره جيفري ساكس وآخرون	نسبة البيانات غير المتوفرة حسب (2) التقرير العربي للتنمية المستدامة (الاسكوا) 2024
مصر	%54.3	%36.9

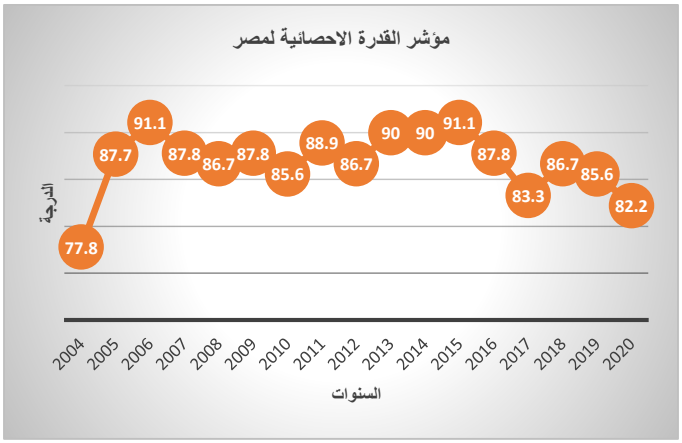
تقييم الجودة (قياس القدرة الإحصائية)

يمكن تعريف القدرة الإحصائية الوطنية على أنها قدرة النظام الإحصائي الوطني ومنظماته وأفراده على جمع وإنتاج وتحليل ونشر إحصاءات وبيانات بجودة عالية وموثوقية لتلبية احتياجات المستخدمين (Paris21,2018) . يعتبر قياس جودة البيانات الإحصائية أمرًا معقدًا: والسبب في ذلك أن العوامل والخصائص التي تشكل الجودة لا يمكن قياسها بشكل مباشر أو ملاحظتها إلا من خلال مؤشرات تقريبية. بالرغم من أن ضمان جودة البيانات يقع في الأساس على الأجهزة الإحصائية، فإن تقييم جودة البيانات يمكن أن يتم من خلال المستخدمين سواء أكانوا مؤسسات وطنية أم مؤسسات دولية متخصصة مثل البنك الدولي وصندوق النقد الدولي، أم منظمات الأمم المتحدة، أم منظمات دولية غير حكومية:

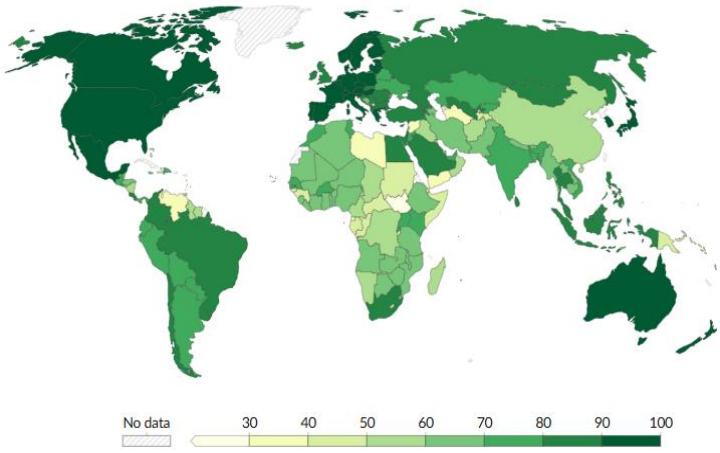
- **صندوق النقد الدولي (معاييرنشر البيانات):** تشكل مبادرات صندوق النقد الدولي في تقييم القدرات الإحصائية تجربة مهمة في مجال المتابعة وتقييم السياسات الوطنية واقتراح سياسات الإصلاح الاقتصادي والمالي بما تتطلبه من بيانات دقيقه وذات جودة ومصداقية عالية. وسعيًا من الصندوق لبلوغ هذه الأهداف فإنه أسس ثلاثة مستويات خاصة بنشر البيانات الاقتصادية والمالية من طرف الدول الأعضاء وبشكل طوعي. وتصف هذه المبادرات المتطلبات المحددة التي يجب أن تلتزم بها البلدان في نشر بياناتها والتي تتعلق بتغطية الإحصاءات المنتجة ودوريتها، ومتطلبات نشر الإحصائيات في أوقات محددة مسبقًا وفقًا لتقويم الإصدار الإحصائي، ومتطلبات البيانات الوصفية. وتم إنشاء المعيار الخاص لنشر البيانات عام 1996، مصر مشترك منذ (2005).

- البنك الدولي (مؤشر القدرة الإحصائية - مؤشر الأداء الإحصائي)

الشكل رقم (2): مؤشر القدرة الإحصائية لمصر للفترة 2004- 2020

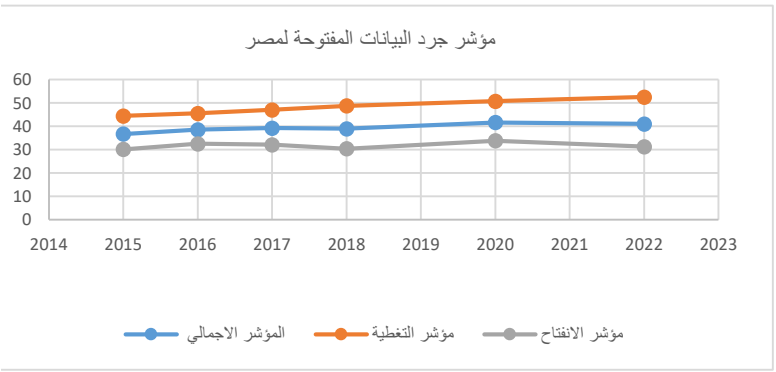


الشكل رقم (3): مؤشر الأداء الإحصائي لدول للعالم عام 2023، مصر مصفنة في الخميس الخامس،ترتيب متقدم جدا عن كثير من الدول النامية.



- مرصد البيانات المفتوحة (مؤشر جرد البيانات المفتوحة)

الشكل رقم (4): مؤشر جرد البيانات المفتوحة لمصر للسنوات 2015-2022



يقوم مؤشر جرد البيانات المفتوحة الذي يعده الذي يعده مرصد البيانات المفتوحة .(Open data watch)بتقييم مدي تغطية (شمول) وإتاحة (انفتاح) الإحصاءات الرسمية بغرض تقييم الفجوات وتعزيز سياسات البيانات المفتوحة وتشجيع الحوار بين المكاتب الإحصائية الوطنية (NSOs)ومستخدمي البيانات. ويقيس هذا المؤشر مدى اكتمال الإحصاءات الوطنية في أي بلد وما إذا كانت بياناتها تلبي المعايير الدولية للانفتاح. يجب أن تكون البيانات التي يتم تقييمها في (ODIN) هي بيانات الدولة الرسمية المنشورة على الموقع الإلكتروني لمكتب الإحصاء الوطني(NSO) أو أي موقع رسمي آخر مرتبط بموقع مكاتب الإحصاء الرسمية.تعتمد درجات الإتاحة والانفتاح على ما إذا كان من الممكن تحميل البيانات بأشكال يمكن قراءتها آليًا وبدون تكاليف معتبرة، وتكون مصحوبة ببيانات تعريفية. تقييما لأداء مصر في جرد البيانات المفتوحة : تقع في الخميس الثالث بدرجة 41%.

جودة البيانات والنمذجة الاقتصادية

يمثّل توفير البيانات ذات الجودة العالية والأعداد الجيّد ضرورة ومتطلب أساسي من متطلبات بناء النماذج الاقتصادية القياسية الموثوقة، وتشكّل العمود الفقري للتحليل الاقتصادي القياسي. وبدون بيانات دقيقة، تصبح النماذج القياسية الأكثر صحة من الناحية النظرية عُرضة للخروج باستنتاجات خاطئة. ويُعدّ جمع البيانات وتجهيزها خطوة حاسمة في خطوات بناء وصياغة النماذج الاقتصادية القياسية.

الشكل رقم (5): مستوى توفر البيانات للنماذج القياسية الاقتصادية الكلية

المصدر تقرير التنمية العربي: الإصدار الثامن

- مستوى توفر البيانات للنماذج القياسية الاقتصادية الكلية لمصر

إجمالا مستوى توفر البيانات في مصر لبناء النماذج متوسط لمرتفع ، يمكن بناء نماذج متوسطة، فجوات معتبرة في البيانات (54%) ، عدم توفر البيانات الفصلية بالشكل المطلوب.

حوكمة البيانات في مصر

تشير إلى مجموعة من العمليات والسياسات التي تهدف إلى ضمان إدارة البيانات بشكل منظم وفعال داخل المؤسسات والحكومات .تركز على توفير بيئة مناسبة لإدارة البيانات من خلال إطار تنظيمي وتشريعي يضمن الشفافية، الأمان، والمسؤولية.

- أطر حوكمة البيانات الحكومية في مصر

الشكل رقم (6) يوضح أطر حوكمة البيانات في مصر

سبل تحسين جودة البيانات في مصر

- تحسين جودة البيانات في مصر يتطلب نهجًا شاملاً يشمل

إنشاء هيئة موحدة لحوكمة البيانات, تحديث وتفعيل الاستراتيجية الوطنية للإحصاء.

- تشكيل هيئة وطنية لحوكمة البيانات تعمل على توحيد المعايير والمنهجيات بين مختلف المؤسسات الإحصائية والوزارات.
- استراتيجية وطنية للإحصاء تتماشي مع التحول الرقمي والذكاء الاصطناعي.
- ضمان التنسيق بين البنك المركزي، الجهاز المركزي للإحصاء، وزارة المالية، ووزارة التخطيط لتحقيق اتساق البيانات الاقتصادية.
- وضع إطار قانوني وتنظيمي لضمان جودة وموثوقية البيانات ومنع أي تلاعب أو تضارب في الأرقام.
- أهمية إصدار قانون لتداول المعلومات: لتعزيز الشفافية والمساءلة.

تعزيز الشفافية و إتاحة البيانات المفتوحة

- ✓ تطوير البوابة الوطنية للبيانات المفتوحة للتسهيل علي للباحثين وصناع القرار الوصول إلى إحصاءات محدثة مجانًا لدعم البحث العلمي وصنع القرار المبني على الأدلة.
- ✓ نشر البيانات وفق جداول زمنية منتظمة، مع الإفصاح عن منهجيات الحساب لضمان المصداقية.
- ✓ إنشاء مرصد وطني مستقل لمراقبة جودة البيانات، وإشراك المجتمع الأكاديمي ومنظمات المجتمع المدني في التقييم الدوري للبيانات الرسمية.

تطوير منظومة الحسابات القومية

- ✓ تحديث أساليب التقدير الحسابية أو استكمال مكونات المنظومة.
- ✓ الاهتمام بكل الأبعاد القطاعية والمكانية والبيئية.
- ✓ تحسين أساليب معالجة إنتاج و عمالة القطاع غير الرسمي.
- ✓ إعداد الحسابات القومية ربع السنوية للناتج المحلي الإجمالي سواء وفق الإنتاج أو الإنفاق أو الدخل.
- ✓ إصدار حسابات فرعية لتسليط الضوء علي المعاملات الاقتصادية لقطاعات معينة بشكل أكثر تفصيلاً مثل الزراعة والسياحة.
- ✓ تقدير الناتج المحلي الإجمالي علي مستوى الأقاليم الجغرافية المختلفة في إطار الحسابات الإقليمية.
- ✓ تطوير نظام المحاسبة البيئية الاقتصادية.

تحسين جودة البيانات في المسوح الإحصائية

- ✓ **اعتماد المعايير والمنهجيات الإحصائية الدولية لضمان اتساق البيانات وقابليتها للمقارنة.**

✓ **التكنولوجيا في جمع البيانات**

- استخدام **أجهزة إلكترونية (CAPI)** بدلاً من الاستمارات الورقية لتقليل الأخطاء وتحسين سرعة المعالجة.
- تضمين **قواعد تحقق آلية (Validation Rules)** لتقليل القيم المتطرفة أو غير المنطقية أثناء الإدخال.

✓ **ضمان الجودة أثناء التنفيذ**

- تطبيق **زيارات مراقبة ميدانية** من قبل مشرفين للتحقق من التزام الباحثين بالإجراءات.
- إجراء **إعادة مقابلات (Re-interviews)** على نسبة من العينة للتحقق من صحة البيانات.

✓ **التنسيق مع السجلات الإدارية**

- الدمج بين المسوح والسجلات الإدارية (مثل بيانات التأمينات الاجتماعية أو القوى العاملة الرسمية) لتحسين الدقة وتقليل العبء على المستجيب.

✓ التحول إلى الرقمنة الكاملة للسجلات الإدارية في الوزارات والهيئات الحكومية لإستخدامها في التحليل الإحصائي.

✓ إطلاق مبادرات تدريبية للكوادر الإحصائية لتحسين وتطوير مهارات جمع البيانات وتحليلها واعداد التقارير.

✓ تنفيذ مبادرات بناء القدرات وبرامج المساعدة الفنية التي تقودها المنظمات الاقليمية مثل الاسكوا، صندوق النقد الدولي ،البنك الدولي.

الاستفادة من التكنولوجيا : الذكاء الاصطناعي، (AI) ، الاعتماد علي مصادر البيانات الحديثة للنمذجة

الشكل رقم (7) : بعض مصادر البيانات غير التقليدية المتاحة للنمذجة القياسية

زيادة الاعتماد على التكنولوجيا، الذكاء الاصطناعي، ومصادر البيانات غير التقليدية يمكن أن تساهم بدرجة كبيرة جدًا في تحسين جودة البيانات وتعزيز النمذجة في مصر، وذلك من خلال عدة مسارات

أولاً :تحسين جودة البيانات

1. **الدقة والاتساق:**

- استخدام أدوات الذكاء الاصطناعي والتعلم الآلي يمكن أن يساعد في اكتشاف الأخطاء، التكرارات، أو القيم الشاذة بسرعة أكبر مما هو ممكن يدويًا
- تقنيات مثل معالجة اللغة الطبيعية (NLP) يمكن أن تُستخدم لتحليل وتوحيد البيانات النصية غير المنظمة.

- تقنيات الذكاء الاصطناعي يمكنها: تحليل البيانات التاريخية واكتشاف الأنماط، مما يُمكّن المؤسسات من التنبؤ بالاتجاهات المستقبلية واتخاذ قرارات استراتيجية مستنيرة.
- التعرف على الصور والفيديو: تُستخدم تقنيات الرؤية الحاسوبية لتحليل البيانات البصرية، مثل أنظمة المراقبة الذكية، لدعم السياسات العامة وتعزيز الأمن.
- تحليل بيانات المرور : 🛢 يُساهم الذكاء الاصطناعي في تحليل بيانات النقل لتخفيف الازدحام وتطوير البنية التحتية، مما يدعم مبادرات المدن الذكية.
- مراقبة جودة الخدمات الحكومية: تُشير الدراسات إلى أن تطبيق تقنيات الذكاء الاصطناعي في تحليل البيانات الضخمة يُحسّن من جودة التقارير المالية، (أثر تحليلات البيانات الضخمة بإستخدام تقنيات الذكاء الإصطناعي على تحسين جودة التقارير

2. **الزمن الحقيقي:(Real-Time)**

- الاعتماد على أجهزة الاستشعار وإنترنت الأشياء (IoT) ومصادر البيانات غير التقليدية (مثل وسائل التواصل الاجتماعي أو بيانات الهاتف المحمول) يمكن أن يوفر بيانات آنية تساعد على اتخاذ قرارات سريعة ودقيقة.

3. **تقليل الفجوات الزمنية والجغرافية:**

- التكنولوجيا تساعد في جمع البيانات من أماكن نائية أو من فئات سكانية يصعب الوصول إليها بالطرق التقليدية.

4. **أتمتة جمع البيانات وتنظيفها:**

- يقلل من الأخطاء البشرية ويرفع من الكفاءة، خصوصًا عند استخدام أنظمة متقدمة لأرشفة البيانات وتوثيقها.

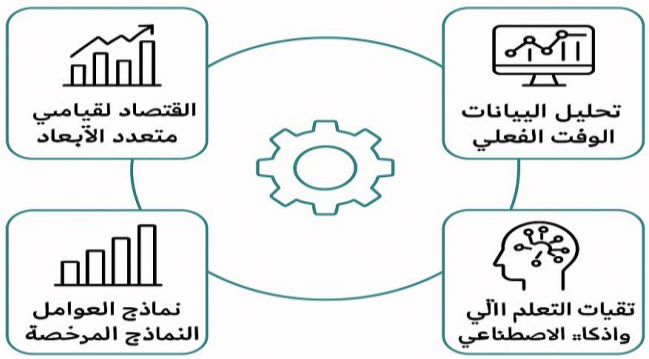
ثانياً: تعزيز النمذجة والتحليل

1. **توفر حجم أكبر وتنوع أعلى من البيانات:**

- المصادر غير التقليدية مثل بيانات الأقمار الصناعية أو وسائل التواصل الاجتماعي يمكن أن تعزز النماذج التنبؤية والاجتماعية والاقتصادية بدقة أعلى.
- يمكن استكمال النماذج القياسية التقليدية التي تعتمد على البيانات الفصلية أو السنوية أو الاستعاضة عنها بنماذج تتضمن **بيانات في الوقت الحقيقي (Real Time Data)** مما يحسّن من مرونة النماذج ويسرّع من استجابة السياسة الاقتصادية للمستجدات الداخلية والخارجية.
- تمكين الباحثين من التحليل التفصيلي العميق: بسبب وفرة البيانات التفصيلية من مصادر مختلفة. وبالتالي تمكين صناع السياسات من فهم الآثار المتفاوتة للسياسات الاقتصادية في المناطق والقطاعات المختلفة.

- التنبؤ الآني(Nowcasting) : استخدام البيانات المنشورة مؤخرًا لتحديث المؤشرات الاقتصادية الرئيسية التي يتم نشرها بتأخر كبير، مثل الناتج المحلي الإجمالي الحقيقي (IMF,2023) . وقد أصبحت ممارسة التنبؤ الآني، أو التنبؤ بالمؤشرات الاقتصادية الحالية أو القريبة جدًا، ذات جدوى عالية بسبب ظهور البيانات في الوقت الفعلي. على سبيل المثال، يمكن للبيانات المالية عالية التكرار وتحليل المشاعر في وسائل التواصل الاجتماعي أن تساعد في التنبؤ بنمو الناتج المحلي الإجمالي قبل نشر الإحصاءات الرسمية.

الشكل رقم (8): تقنيات النمذجة الحديثة



2. **<u>نماذج أكثر تعقيدًا ودقة</u>**

- يمكن للذكاء الاصطناعي تطبيق نماذج متقدمة (مثل الشبكات العصبية أو الغابات العشوائية) التي يصعب تشغيلها على بيانات قليلة أو غير منظمة..

زيادة الاعتماد علي الذكاء الاصطناعي ← مصادر بيانات غير تقليدية ← الاعتماد علي تقنيات النمذجة الحديثة

3. **<u>التكامل بين مصادر البيانات:</u>**

- التكامل بين مصادر البيانات باستخدام تقنيات الربط والدمج الذكي للبيانات، يمكن بناء صورة أكثر شمولية تساعد في فهم العلاقات المعقدة بين الظواهر المختلفة

ثالثاً: تحسين عمليات اتخاذ القرار

- **توفر بيانات دقيقة وفي توقيت مناسب تُحسن من دعم القرار في السياسات العامة، التخطيط الاقتصادي، وإدارة الأزمات مثل الأوبئة أو الكوارث الطبيعية.**

دمج التكنولوجيا والذكاء الاصطناعي والمصادر غير التقليدية هو ركيزة أساسية لتحسين جودة البيانات في مصر، لكن النجاح في ذلك يتطلب تخطيطًا استراتيجيًا، استثمارًا في الكفاءات، وتعاونًا مؤسسيًا فعالًا.