دراسة تطبيقات البيانات الكبيرة "Big Data" في الدول العربية

د. سومية لطفي

الدائرة الاقتصادية والفنية
صندوق النقد العربي

(قدمت الورقة لاجتماع مبادرة "عربستات")

أكتوبر 2018

صندوق النقد العربي 2018

حقوق الطبع محفوظة

الآراء الواردة في هذه الدراسة لا تمثل بالضرورة وجهة نظر صندوق النقد العربي، وتبقى معبرة عن وجهات نظر مؤلف الدراسة.

لا يجوز نسخ أو اقتباس أي جزء من هذه الدراسة أو ترجمتها او إعادة طباعتها بأي صورة دون موافقة خطية من صندوق النقد العربي إلا في حالات الاقتباس القصير بغرض النقد والتحليل مع وجوب ذكر المصدر.

توجه جميع المراسلات على العنوان التالي:

الدائرة الاقتصادية والفنية

صندوق النقد العربي

ص.ب 2818، أبو ظبي دولة الامارات العربية المتحدة–

هاتف: 6171552 – 2 – 971+

فاكس: 6326454 – 2 – 971+

البريد الالكتروني: economic@amfad.org.ae

موقع الصندوق الالكتروني: www.amf.org.ae

قائمة المحتويات

1. الملخص التنفيذي.....4

2. مفهوم البيانات الكبيرة والعوامل المؤثرة فيها.....7

3. أهمية البيانات الكبيرة للسلطات العربية.....11

4. غرض ونتائج استبيان تطبيقات البيانات الكبيرة في الدول العربية.....12

5. أفضل الممارسات لتطبيقات البيانات الكبيرة في الإحصاءات الرسمية ولأغراض الاقتصاد الكلي.....33

6. مقترحات لوضع إطار لبرنامج عمل إقليمي لدعم تطبيق البيانات الكبيرة في الدول العربية.....41

1.6 إنشاء فريق عمل إقليمي معني بالبيانات الكبيرة.....41

2.6 مبادرة إقليمية لبناء القدرات في مجال البيانات الكبيرة.....44

3.6 إنشاء شبكة إقليمية للتعاون وبناء الشراكات في مجال البيانات الكبيرة..... 45

7. الخلاصة والتوصيات.....46

8. المرفقات.....49

1.8 استبيان تطبيقات البيانات الكبيرة في الدول العربية.....49

2.8 المراجع.....59

3.8 الأشكال والجداول.....61

1. الملخص التنفيذي

تهدف دراسة تطبيقات البيانات الكبيرة في الدول العربية إلى التعرف على مدى الوعي والاهتمام بمنهجيات ونظم وقوانين إدارة البيانات الكبيرة، التي عُرفت بالبيانات التي تحتاج لأدوات وتقنيات غير تقليدية لتحليلها والاستفادة منها نظراً لكبر حجمها وتنوعها، ذلك بالاعتماد على نتائج استبيان صندوق النقد العربي حول تطبيقات البيانات الكبيرة في الدول العربية، الذي يهدف إلى تقييم الوضع الراهن ومدى جاهزية السلطات العربية في مجال استخدام البيانات الكبيرة، وكذلك التعرف على احتياجاتها من مصادر البيانات وطرق تحليل البيانات والمهارات اللازمة لذلك. شملت الدراسة الجهات المعنية في وزارات المالية، والمصارف المركزية ومؤسسات النقد العربية والأجهزة الإحصائية في الدول العربية. من جهة أخرى، تضمنت الدراسة استعراضاً عن بعض الممارسات في مجال البيانات الكبيرة لأغراض الإحصاءات الرسمية وتحليل الاقتصاد الكلي، ومقترحات لوضع إطار لبرنامج عمل إقليمي لدعم تطبيق البيانات الكبيرة في الدول العربية، يشمل إطلاق مبادرة إقليمية لبناء القدرات في مجال البيانات الكبيرة واستراتيجية بناء القدرات بهدف معالجة القضايا المتعلقة بالمنهجية وجودة البيانات والتقنيات والوصول إلى البيانات والتشريعات والسياسات المرتبطة بسرية وخصوصية البيانات. كذلك توفير طرق التحليل المناسبة التكلفة بشأن استخدام البيانات الكبيرة في الدول العربية. كما يشمل المقترح إنشاء شبكة إقليمية للتعاون وتبادل الخبرات، وتعزيز التعاون مع جهات إقليمية ودولية معنية بالبيانات الكبيرة.

تناول استبيان صندوق النقد العربي جمع معلومات أساسية من جهات رسمية حول تطبيقات البيانات الكبيرة في الدول العربية، بما فيها السياق المؤسسي التنظيمي، ورصد المعلومات حول المشاريع االراهنة والمستقبلية المرتبطة بتطبيقات البيانات الكبيرة. تم التركيز في الاستبيان على مصادر البيانات والأدوات المستخدمة من قبل الجهات العربية المعنية، إضافة إلى الشراكات التي تقوم بها مع مزودي البيانات والتعاون مع المؤسسات الإقليمية والدولية والتحديات التي تواجهها في إنجاز وتطوير مشاريع البيانات الكبيرة.

تبيَن من الردود التي تم استلامها من قبل الصندوق مدى اهتمام الدول العربية بتطبيقات البيانات الكبيرة، حيث استجابت سبعة عشر جهة للاستبيان، شملت وزارات المالية، والمصارف المركزية ومؤسسات النقد العربية، ومراكز الإحصاء العربية. تُمثل هذه الجهات ثلاث عشر دولة عربية هي: المملكة الأردنية الهاشمية، دولة الإمارات العربية المتحدة، مملكة البحرين، الجمهورية التونسية، المملكة العربية السعودية، سلطنة عُمان، دولة فلسطين، دولة الكويت، جمهورية لبنان، دولة ليبيا، المملكة المغربية، جمهورية مصر العربية، والجمهورية الإسلامية الموريتانية. بلغت نسبة الردود على الاستبيان 6% من وزارات المالية، و41% من البنوك المركزية ومؤسسات النقد العربية، و53% من أجهزة الإحصاء في الدول العربية.

فيما يلي الاستنتاجات الرئيسية للاستبيان:

إهتمام الإدارات العليا بموضوع البيانات الكبيرة. حوالي 88% من الجهات التي استجابت للاستبيان أشارت إلى إهتمام الإدارة العليا بموضوع البيانات الكبيرة، منها حوالي 41% قام بإعداد خطة أو استراتيجية لاستخدام البيانات الكبيرة، وحوالي 29% قام بتكوين لجنة داخلية تناقش وتستعرض الموضوع.

العمل على مشاريع البيانات الكبيرة: أظهرت نتائج الاستبيان أنه بالرغم من اهتمام الادارات العليا بالبيانات الكبيرة وتطلعاتها بشأنها، حيث ينظر حوالي 82% من الجهات المستطلعة لبدء مشاريع البيانات الكبيرة خلال الفترة المقبلة، إلاّ أن حوالي 47% منها فقط يعمل فعلياً على مشاريع للبيانات الكبيرة. كما يلاحظ من الإجابات أن حوالي 34% من المشاريع تمت الموافقة عليها ولم يتم تمويلها بعد، في حين حوالي 13% فقط اكتملت.

استخدام مصادر البيانات الكبيرة يواجه تحديات عديدة: حوالي 59% من الجهات التي أجابت على الاستبيان لا تستخدم حالياً أي مصادر للبيانات الكبيرة، وحوالي 47% أشارت إلى عدم توفر إمكانية الوصول المباشر إلى مصادر البيانات الكبيرة والحصول عليها بصفة دورية. تشمل التحديات كذلك الحاجة إلى المهارات التقنية والخبرات الوطنية المختصة في البيانات الكبيرة، وكذلك الحاجة إلى معالجة قضايا التشريعات المتعلقة بسريَة وخصوصية البيانات، والتعاون مع الجهات المنتجة للبيانات وإبرام اتفاقيات تبادل البيانات، إلى جانب توفير الموارد المالية الكافية لتنفيذ مشروعات البيانات الكبيرة وضمان استمراريتها.

تحليل البيانات الكبيرة محدود حالياً: بينت الإجابات على الاستبيان أن نسبة ملحوظة 75% لا تزال في بداية مراحل الإستعداد، وتعتبر غير متمكنة حالياً لبدء عمليات تحليل البيانات الكبيرة أو الحصول على المخرجات المتوقعة من استخدامها بشكل منتظم. كما أن حوالي 69% من الجهات المستطلعة لا تتوفر على وسائل يُمكن استخدامها في تحليل البيانات الكبيرة. وحوالي 22% فقط من الجهات المستطلعة أفاد أن العمل جارٍ على تطوير أدوات تحليل البيانات الكبيرة داخلياً. يتمثل أحد التحديات الرئيسة في عدم توفر الموارد البشرية الملائمة وتقنيات المعلومات اللازمة.

مجالات استخدام البيانات الكبيرة: تتصدر الإحصاءات الاقتصادية والمالية القائمة، بحوالي 46% من مشاريع البيانات الكبيرة المحتملة، و38% لأغراض إحصاءات السياحة، و31% لإحصاءات الأسعار، و31% لإحصاءات النقل، و31% للإحصاءات الديموغرافية والاجتماعية.

توظيف خبراء مختصين وتدريب الموظفين في مجال البيانات الكبيرة محدود حالياً: عبرت الجهات المعنية عن حاجتها إلى مهارات وخبرات في مجال البيانات الكبيرة، حيث أن حوالي 35% منها يقوم بتوظيف خبراء مختصين وتدريب الموظفين لديها. يتمثل النقص خصوصاً في مهارات التحليل والقدرة على الإبداع والإبتكار

وإتقان علوم البيانات والذكاء الاصطناعي وتنقيب البيانات (data mining) وطرق التنبؤ، كذلك استخدام أنظمة/آليات تقنية المعلومات والبرمجة، وتجميع ونقل وتخزين ومعالجة البيانات.

تعاون الجهات المعنية المستطلعة مع نظرائها في مجال البيانات الكبيرة: ترغب الجهات المستطلعة في التعاون مع نظرائها في مجال البيانات الكبيرة، والمشاركة مع دول ومؤسسات أخرى أو لجان أو ورش عمل، مما يساعدها في عملية استكشاف البيانات الكبيرة وتحليلها، الأمر الذي تعتبره أغلبية الجهات المعنية مهمة معقدة ومتعددة الأوجه.

المساهمة في وضع خارطة طريق /أو خطة للتعاون الإقليمي حول استخدام وتطوير البيانات الكبيرة: عبرت كل الجهات المستطلعة عن استعدادها للمساهمة في وضع خارطة طريق و/أو خطة للتعاون الإقليمي حول استخدام وتطوير البيانات الكبيرة وإدارة التحديات المرتبطة بها.

المجالات ذات الأولوية: حددت الجهات العربية المعنية عدداً من المجالات ذات أولوية للتركيز عليها، تغطي مجموعة واسعة من المواضيع الإحصائية البحتة منها: جودة البيانات (تمثل 100% من اهتمام الجهات المُستطلعة)، تليها الإحصاءات والمؤشرات (86%) وتقنيات أخذ العينات (73%). كما أن حوالي 67% من الجهات يهتم بالتعرف على مصادر البيانات الكبيرة. يعكس ذلك، التحديات الناجمة عن الاستخدامات الجديدة لمصادر البيانات الكبيرة.

قضايا السرية والخصوصية فيما يتعلق بمصادر البيانات الكبيرة وطرق معالجتها: حوالي 53% من الجهات المستطلعة عبَرت عن عدم وجود إطار قانوني لديها لمعالجة قضايا سرية وخصوصية البيانات. كما رأت أن التشريعات القانونية تمثل أحد التحديات الجوهرية التي تواجه مجال البيانات الكبيرة.

أهم مجالات مشاريع البيانات الكبيرة الحالية والمستقبلية: هناك 41% من مشاريع البيانات الكبيرة تتعلق بإعداد وإصدار الاحصاءات، و35% مشاريع تجريبية "Pilots"، بينما 24% منها هي مشاريع استكشافية للبحث.

التعاون والشراكات مع مزودي البيانات الكبيرة محدودة حالياً: حوالي 59% من الجهات المستطلعة أشارت إلى عدم وجود تعاون أو شراكات في مجال البيانات الكبيرة مع أي منظمات أو موفري بيانات. بالنسبة للجهات الأخرى، حوالي 67% منها أشارت إلى أن الشراكات الحالية في مجال البيانات الكبيرة في مرحلة المناقشة، وحوالي 17% تم عقد اتفاقيات حولها، في حين 16% فقط تم اعتمادها. فيما يخص نوعية المؤسسات الشريكة في مجال البيانات الكبيرة، بيَنت الإجابات أن 45% من الشراكات الحالية هي شراكات مع مؤسسات دولية حكومية، و22% مع مؤسسات تجارية غير حكومية، و11% مع المؤسسات الأكاديمية .

بهدف تطوير العمل في مجال البيانات الكبيرة، وتعزيز دور الصندوق كمركز للتشاور والوعي بأهمية تطبيقات البيانات الكبيرة، تم وضع مقترح لبرنامج عمل إقليمي لدعم تطبيق البيانات الكبيرة في الدول العربية، يشمل ثلاث محاور رئيسية:

المحور الأول: إمكانية إنشاء فريق عمل إقليمي معني بالبيانات الكبيرة أو الإستفادة من الأطر القائمة حالياً.

المحور الثاني: مبادرة إقليمية لبناء القدرات في مجال البيانات الكبيرة.

المحور الثالث: إنشاء شبكة إقليمية للتعاون وبناء الشراكات في مجال البيانات الكبيرة.

2. مفهوم البيانات الكبيرة والعوامل المؤثرة فيها

تشكل ثورة البيانات تحدي للتوجه التقليدي ليس فقط فيما يتعلق بجمع البيانات وتحليلها، بل أيضاً من حيث التقنية والمهارات الجديدة التي تتطلبها، والاستراتيجيات التي تضعها الدول للاستفادة منها، نظراً لمدى الأهمية الاستراتيجية للبيانات الكبيرة للعديد من القطاعات الاقتصادية والمالية وسواها.

يتزايد حجم ونوع البيانات المتوافرة بشكل كبير، وأصبح يطلق عليها مسمى "البيانات الكبيرة". تتألف البيانات الكبيرة من البيانات المهيكلة (structured)، وهي البيانات المنظمة في صورة جداول أو قواعد بيانات تمهيداً لمعالجتها، والبيانات غير المهيكلة (non-structured)، التي تشكل النسبة الأكبر من البيانات، وهي البيانات التي يولدها الأشخاص يومياً من كتابات نصية وصور وفيديو ورسائل ونقرات على مواقع الانترنت... إلخ. تعتبر البيانات كبيرة إذا توفرت فيها ثلاثة عوامل رئيسة: الحجم - وهو عدد "التيرابايت" (Terabytes) من البيانات، التنوع - وهو تنوع هذه البيانات ما بين مهيكلة وغير مهيكلة، والسرعة - وهو مدى سرعة تواتر حدوث البيانات.

لعل الخدمات المالية والمصرفية أحد القطاعات التي تزدهر بتوافر البيانات الكبيرة نظراً لحجم المعاملات التي يتم إجراءها يومياً في جميع أنحاء العالم. كما أن التطورات التقنية اليوم، إلى جانب توافر كمية كبيرة من المعلومات، تجعل البيانات الكبيرة ذات أهمية متزايدة للقطاع المالي والمصرفي، كذلك للأجهزة الإحصائية وغيرها في السنوات المقبلة، كما تثير تحديات جديدة لهذه الجهات. بدأت الأجهزة الإحصائية والمصارف المركزية في التفكير الاستراتيجي العميق إزاء البيانات الكبيرة، إدراكاً بأهمية استخدامها في الإحصاءات الرسمية وطرق التحليل.

مع تزايد اهتمام مختلف المؤسسات بتخزين ومعالجة واستخلاص القيمة من البيانات بجميع أشكالها وأحجامها، تستمر الأنظمة التي تدعم كميات كبيرة من البيانات في التزايد، كما يتزايد طلب السوق على المنصات التي تساعد على تنظيم وأمن البيانات الكبيرة، وتمكين المستخدمين النهائيين من تحليل تلك البيانات. يتجلى التحدي في كيفية استخدام البيانات الكبيرة في التطبيقات العملية، وحجم الفرص التي تمثلها أمام المصارف المركزية والأجهزة الإحصائية. هناك العديد من التساؤلات التي تطرح في هذا الشأن:

كيف يمكن للمصارف المركزية والأجهزة الإحصائية الرسمية تحقيق أفضل استفادة ممكنة من البيانات الكبيرة؟ وكيف يمكن أن تساعد البيانات الكبيرة المصارف المركزية على اتخاذ تدابير سياساتها العامة (policy measures) في الوقت المناسب؟ وكيف يمكن للأجهزة الإحصائية الرسمية العربية معالجة القضايا المتعلقة بمنهجية جودة وتقنية البيانات بشكل كاف، وخلق بيئة تعزز ثقة العموم في استخدام البيانات الكبيرة للإحصاءات الرسمية؟

1.2 تعريف البيانات الكبيرة

لا يوجد تعريف دقيق للبيانات الكبيرة، حيث توصف بشكل عام أنها بيانات متعددة الأنواع والمصادر والأحجام. يُعرف الخبراء البيانات الكبيرة أنها مجموعة البيانات ذات حجم يفوق قدرة معالجتها باستخدام أدوات قواعد البيانات التقليدية من تجميع، وتقاسم ونقل، وتخزين، وإدارة وتحليل تلك البيانات في غضون فترة زمنية مقبولة. أما من وجهة نظر مقدمي الخدمات، فهي تشمل الأدوات والعمليات التي تحتاجها المنظمات للتعامل مع كمية كبيرة من البيانات لأغراض التحليل.

في نفس السياق، تصف المنظمة الدولية للمعايير (ISO) البيانات الكبيرة أنها *"مجموعة أو مجموعات من البيانات التي لها خصائصها الفريدة (مثل الحجم، السرعة، التنوع، التباين، صحة البيانات... إلخ)، ولا يمكن معالجتها بكفاءة باستخدام التكنولوجيا الحالية والتقليدية لتحقيق الاستفادة منها"*. بينما حسب تعريف الاتحاد الدولي للاتصالات (ITU)، يشير مصطلح البيانات الكبيرة إلى مجموعات البيانات التي تتميز بكونها فائقة حجماً وسرعة أو تنوعاً، بالقياس إلى أنواع مجموعات البيانات المعهودة الاستخدام".

2.2 تصنيف البيانات الكبيرة

كثير منا يعتقد أن البيانات الكبيرة تصنف وفقاً للحجم فقط. في حقيقة الأمر، يمكن تصنيف البيانات الكبيرة وفقاً لما يطلق عليه مبدأ "V’s3"، الذي يشمل الحجم (volume)، والتنوع (variety)، والسرعة (velocity). نظراً للفرص والتحديات التي تواجهها المؤسسات عند دمج البيانات الكبيرة في عملياتها القائمة، تم توسيع تعريف البيانات الكبيرة وفقاً لمبدء (5V's) . يشمل مبدأ (V's5) مايلي:

الحجم (Volume): حجم البيانات المستخرجة من مصدر ما، وهو ما يحدد قيمة وإمكانات البيانات لكي تصنف ضمن البيانات الكبيرة أو الضخمة؛ وقد يكون من الخصائص الأكثر أهمية في تحليل البيانات الكبيرة. كما أن وصف البيانات بالضخمة لا يحدد كمية معينة؛ فالحجم يقاس عادة ب "Yottabyte" أو"Zettabyte" أو"Exabyte" أو"Petabytes" أو"Terabyte" "Gigabyte" أو"Megabyte" أو"Kilobyte" أو"Byte". يقدر أن 90% من البيانات الموجودة في العالم اليوم قد استحدثت خلال السنتين الأخيرتين، بواسطة أجهزة. يبين الجدول 1 قياسات حجم البيانات.

جدول 1. قياسات حجم البيانات

القياس	القيمة	البايت "bytes"
Byte (B)	1	1
Kilobyte (KB)	$1,024^1$	1,024
Megabyte (MB)	$1,024^2$	1,048,576
Gigabyte (GB)	$1,024^3$	1,073,741,824
Terabyte (TB)	$1,024^4$	1,099,511,627,776
Petabyte (PB)	$1,024^5$	1,125,899,906,842,624
Exabyte (EB)	$1,024^6$	1,152,921,504,606,846,976
Zettabyte (ZB)	$1,024^7$	1,180,591,620,717,411,303,424
Yottabyte (YB)	$1,024^8$	1,208,925,819,614,629,174,706,176

التنوع (Variety): يقصد بها تنوع البيانات المستخرجة، والتي تساعد المستخدمين، باحثين أو محللين على اختيار البيانات المناسبة لمجال بحثهم وتتضمن بيانات مهيكلة في قواعد بيانات وبيانات غير مهيكلة تأتي من طابعها غير الممنهج، مثل: الصور ومقاطع وتسجيلات الصوت وأشرطة الفيديو والرسائل القصيرة وسجلات المكالمات وبيانات الخرائط (GPS) وغيرها، وتتطلب وقتاً وجهداً لتهيئتها في شكل مناسب للتجهيز والتحليل.

السرعة (Velocity): يقصد بها سرعة إنتاج واستخراج البيانات لتغطية الطلب عليها، حيث تعتبر السرعة عنصراً حاسماً في اتخاذ القرار بناء على هذه البيانات، وهو الوقت الذي نستغرقه من لحظة وصول هذه البيانات إلى لحظة الخروج بالقرار بناء عليها. سابقاً كانت الشركات تستخدم لمعالجة مجموعة صغيرة من البيانات المخزنة في صورة بيانات مهيكلة في قواعد بيانات عملية تسمى "Batch Process"، حيث كان يتم تحليل كل مجموعة بيانات الواحدة تلو الأخرى في انتظار وصول النتائج. مع ازدياد حجم البيانات وسرعة تواترها أصبحت الحاجة أكثر إلحاحاً إلى نظام يضمن سرعة فائقة في تحليل البيانات الكبيرة في الوقت الآني "Real Time" أو شبه الآني، وأدت تلك الحاجة إلى ابتكار تقنيات وحلول مثل (Apache)، و(SAP HANA) و(Hadoop)، وغيرها من التقنيات.

المصداقية (Veracity): تشير المصداقية إلى التحيز في البيانات باعتبارها أحد أكبر التحديات لتحقيق قيمة وجودة البيانات الكبيرة.

التقلب (Volatility): يمكن أن يؤدي تغيير التقنية أو بيئات الأعمال التي تولد البيانات الكبيرة، إلى التأثير على نتائج تحليل البيانات الكبيرة، فضلاً عن هشاشة البيانات الكبيرة كمصدر للبيانات.

3.2 مصادر البيانات الكبيرة

يمكن تصنيف البيانات الكبيرة حسب مصادرها على النحو التالي:

جدول 2. بعض مصادر البيانات الكبيرة

- المصادر الناشئة عن إدارة أحد البرامج، سواء كان برنامجاً حكومياً أو غير حكومي، كسجلات التأمين والسجلات المصرفية، والسجلات الطبية الإلكترونية وزيارات المستشفيات وغيرها.
- المصادر التجارية أو ذات الصلة بالمعاملات الناشئة عن معاملات بين كيانين، على سبيل المثال: معاملات البطاقات الائتمانية والمعاملات التي تجرى عن طريق الإنترنت بوسائل مثل الأجهزة المحمولة.
- مصادر شبكات أجهزة الاستشعار، على سبيل المثال: التصوير بالأقمار الصناعية، وأجهزة إستشعار الطرق، وأجهزة إستشعار المناخ.
- مصادر أجهزة التتبع، على سبيل المثال: تتبع البيانات المستمدة من الهواتف المحمولة والنظام العالمي لتحديد المواقع.
- مصادر البيانات السلوكية، على سبيل المثال، عدد مرات البحث على الإنترنت عن معلومات مثل منتج أو خدمة أو أي نوع آخر من المعلومات، وعدد مرات مشاهدة إحدى الصفحات على الإنترنت.
- مصادر البيانات المتعلقة بالآراء، على سبيل المثال، التعليقات على وسائط التواصل الاجتماعي.

4.2 النمو السريع للبيانات الكبيرة/البيانات غير المهيكلة

يُبين الجدول رقم (3) بعض الأمثلة من الواقع حول النمو السريع لاستخدامات البيانات الكبيرة.

جدول 3. أمثلة من الواقع حول النمو السريع للبيانات الكبيرة/البيانات غير المهيكلة

يتم تحميل 400 ساعة من الفيديو على "YouTube" كل دقيقة، ويشاهد أكثر من مليار ساعة من مقاطع فيديو "YouTube" يوميًا.
يتم إنشاء 571 موقع إلكتروني جديد في كل دقيقة من اليوم.
تتلقى العلامات التجارية والمنظمات عبر"Facebook" حوالي 34722 "Likes" كل دقيقة من اليوم.
يتم تحميل 100 "تيرابايت" من البيانات يومياً إلى "Facebook".
يتوفر "Whatsapp" على أكثر من مليار مستخدم، ويتم تداول أكثر من 42 مليار رسالة وحوالي 1.6 مليار صورة بشكل يومي.
تتعامل فيسبوك "Facebook" مع أكثر من 50 مليار صورة من مستخدميها.
يتعامل "Google" مع حوالي 100 مليار عملية بحث في الشهر.
سجلات مشتريات السوبر ماركت (Wallmart): تشمل أزيد من 1 مليون معاملة في الساعة.

5.2 العوامل المساعدة على تطور استخدام البيانات الكبيرة

من بين أهم العوامل والتطورات التي تساعد على استخدام البيانات الكبيرة:

- التقدم في تقنية المعلومات التي خفضت عبء جمع البيانات وتخزينها، وتكاليف معالجتها.
- تطوير مصادر جديدة للبيانات وتحسين إمكانية الوصول إلى مجموعات البيانات الكبيرة القائمة.
- التوفر على بيانات ذات حجم ضخم يتم تجميعها لأغراض مثل الأعمال التجارية، والإدارة، والرعاية الصحية، والأرصاد الجوية، وحركة المرور، والتي يمكن استخدامها لمجموعة أغراض غير تلك التي تم تجمعيها من أجلها.
- تطوير موازٍ لطرق وأساليب جديدة مبتكرة وفعالة لاستغلال وتحليل البيانات الكبيرة.
- التقارب بين البيانات الكبيرة، وإنترنت الأشياء، وتحليل ومعالجة معظم البيانات بالقرب من مصدرها، ما يطلق عليه "at the edge".

6.2 مخاوف الخصوصية (Privacy Concerns) حول الاستخدامات الجديدة للبيانات الكبيرة

- **بالنسبة للأفراد**، تتعلق المخاوف بالإفصاح عن البيانات الشخصية الطبية والمالية والقانونية الحساسة وغيرها من البيانات الحساسة، والاستخدامات التي من شأنها أن تؤدي إلى نتائج تمييزية، والاستخدامات لأغراض الضرائب والتحقيقات القانونية وغيرها من الأغراض الحكومية.

- **بالنسبة للشركات،** فإن المخاوف هي الكشف والإفراج للجمهور عن: تسويق القيمة تجارياً ومجموعات البيانات الأخرى، ومعلومات عن الأساليب والمصادر المستخدمة لإنتاج تلك البيانات، والإفصاح للمنافسين عن المعلومات الاستراتيجية الهامة على التسعير والتكاليف والأرباح والأسواق. واستخدام هذه المعلومات لأغراض الضرائب وللأغراض التنظيمية والتحقيقية والقانونية، وغيرها من الأغراض.
- **بالنسبة للقطاع العام،** هناك حاجة أيضاً إلى معالجة شواغل الأفراد والشركات وحماية بياناتهم. ويتعين على القطاع العام التأكد من عدم الكشف عن:
  - الاسم والعنوان أو معلومات التعريف الأخرى التي يمكن استخدامها لأغراض التسويق وغيرها من الأغراض التجارية.
  - التفاصيل الشخصية، بما في ذلك المعلومات مثل الحالة الزوجية والصحية والدخل.
  - أي معلومات يمكن أن يؤدي استخدامها إلى نتائج تمييزية في مجالات مثل العمالة أو الأهلية للحصول على القروض أو الأهلية للبرامج الحكومية.
  - كما أن أي إفصاح للوكالات غير الإحصائية يغير ميزان القوى بين الأفراد والحكومة، بما في ذلك استخدام البيانات التي تم جمعها لأغراض إحصائية ولأغراض الضرائب والتنظيم والتحقيق وغيرها من الأغراض غير الإحصائية.

3. أهمية البيانات الكبيرة للسلطات العربية

هناك فرص لوزارات المالية والمصارف المركزية ومؤسسات النقد العربية والأجهزة الاحصائية في الدول العربية لكي تستفيد من استخدام البيانات كمصدر للرؤية الهادفة وأيضا لتمكين التطور السريع لعمليات تحليل واستخراج هذه البيانات والأنشطة الخاصة بهذه الجهات. لا شك أن المصارف المركزية تركز اهتمامها حول التحدي المتمثل في تحليل مجموعات البيانات الكبيرة، فهناك عدداً متزايداً من المواضيع ذات الأهمية الكبيرة للمصارف المركزية والسلطات الرقابية، مثل التنبؤ والإحصاءات، إضافة إلى تزايد عدد البحوث التي تقوم بها المصارف المركزية، كل ذلك يتطلب تقييماً لكميات هائلة ومتزايدة من البيانات.

إذا أمكن حل التحديات المتعلقة بالمفاهيم والتعاريف والتمثيل (representativeness)، فمن المرجح أن تغير البيانات الكبيرة العمليات الداخلية للمصارف المركزية، بل أيضاً النظم الاقتصادية والمالية الخارجية التي تشرف عليها السلطات. ويمكِن استخدام البيانات الكبيرة في تعزيز وتحسين الفهم للتنمية الاقتصادية، واستجابة صناع القرار للتغيرات بسرعة أكبر. كما يمكن أن يغير كيف تقود الشركات المالية والوكلاء الاقتصاديين أنشطتهم وأعمالهم التجارية. ومع هذه التغيرات، فإنه من المهم بالنسبة للمصارف المركزية فهم كيف يمكن لاستخدام البيانات الكبيرة أن يؤثر على الاستقرار النقدي والمالي، فضلاً عن النمو الاقتصادي وفرص العمل. من جهة أخرى، يمكن للأجهزة الإحصائية العربية الاستفادة من توافر البيانات الكبيرة لتحديث عملها وإثراء مصادر الإحصاءات الرسمية، وتوفير بيانات أكثر تفصيلاً في الوقت

المناسب بتكلفة أقل بكثير من مجموعات المسوح الجديدة أو الموسعة، إضافة إلى تشجيع البحوث حول استخدام البيانات الكبيرة في مجال الإحصاء، بما يساهم في تعزيز العلاقة بين البيانات والتنمية المستدامة.

4. غرض ونتائج استبيان تطبيقات البيانات الكبيرة في الدول العربية

الغرض من استبيان تطبيقات البيانات الكبيرة في الدول العربية

يتمثل الهدف من الاستبيان في التعرف على احتياجات الدول العربية في مجال البيانات الكبيرة وطرق التحليل المتقدمة، ورصد معلومات واسعة حول مشاريع البيانات الكبيرة المحتملة في وزارات المالية والمصارف المركزية ومؤسسات النقد العربية والمجتمع الإحصائي، وكذلك عن الشراكات ومصادر البيانات والأدوات المستخدمة، بما يساعد صندوق النقد العربي بالتعاون مع المؤسسات والأطر الدولية المعنية، لوضع برنامج تنسيق إقليمي يشمل بناء القدرات في مجال البيانات الكبيرة وطرق التحليل المتقدمة، بما يخدم وزارات المالية، والمصارف المركزية ومؤسسات النقد العربية والأجهزة الإحصائية في الدول العربية. كذلك بما يعزز الاهتمام الاستراتيجي وأهداف هذه الجهات لتطوير مشاريع "البيانات الكبيرة". إلى جانب التعرف على الوضع الراهن لاستخدام البيانات الكبيرة والتحديات التي تواجه المؤسسات العربية في إنجاز وتطوير مشاريع البيانات الكبيرة. كما يهدف أيضاً إلى معالجة القيمة الحالية والفرص المستقبلية من البيانات الكبيرة وأدوات تحليل البيانات من قبل المصارف المركزية والسلطات الرقابية المالية والصناعة المالية، وكذلك الأجهزة الإحصائية لتقديم رؤى قيمة لتطوير البيانات الكبيرة. كما يهدف إلى الاستفادة من تجارب الدول التي حققت تقدماً في هذا المجال.

أُرسل الاستبيان إلى المدراء ممثلي وزارات مالية والمصارف المركزية والأجهزة الإحصائية العربية، أعضاء مبادرة "عربستات" بتاريخ **17** أغسطس 2017. استلم الصندوق ردود عددها 17 حول الاستبيان معبأة من قِبل الجهات المعنية في ثلاث عشر دولة عربية هي: المملكة الأردنية الهاشمية، دولة الإمارات العربية المتحدة، مملكة البحرين، الجمهورية التونسية، المملكة العربية السعودية، سلطنة عُمان، دولة فلسطين، دولة الكويت، جمهورية لبنان، دولة ليبيا، المملكة المغربية، جمهورية مصر العربية، والجمهورية الإسلامية الموريتانية.

تجدر الاشارة إلى أنه تم اعتماد التعريف التالي الواسع النطاق لمفهوم البيانات الكبيرة:

"البيانات الكبيرة هي مصدر بيانات ذات حجم مرتفع، ومجموعة متنوعة من البيانات، الأمر الذي يتطلب أدوات وطرق جديدة لتجميعها، وإدارتها، ومعالجتها بطريقة فعالة".

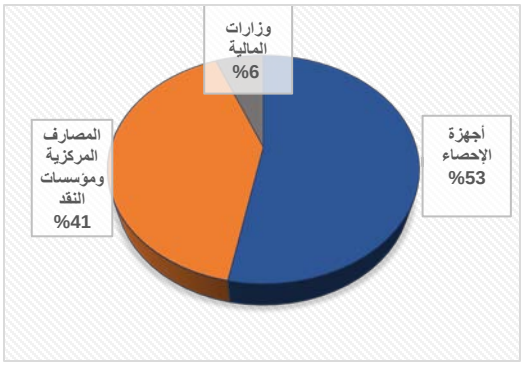
نتائج استبيان تطبيقات البيانات الكبيرة في الدول العربية

الجهات المشاركة في الاستبيان

تم استلام ردود حول الاستبيان من 17 جهة، تشمل وزارات المالية، والمصارف المركزية ومؤسسات النقد العربية، وأجهزة الإحصاء في الدول العربية. تمثل هذه الجهات ثلاث عشر دولة عربية هي: المملكة الأردنية الهاشمية، دولة الإمارات العربية المتحدة، مملكة البحرين، الجمهورية التونسية، المملكة العربية السعودية، سلطنة عُمان، دولة فلسطين، دولة الكويت، جمهورية لبنان، دولة ليبيا، المملكة المغربية، جمهورية مصر العربية، والجمهورية الإسلامية الموريتانية.

بلغت نسبة الردود على الاستبيان 53% من قبل الأجهزة الإحصائية، و41% من البنوك المركزية ومؤسسات النقد العربية، و6% من وزارات المالية في الدول العربية.

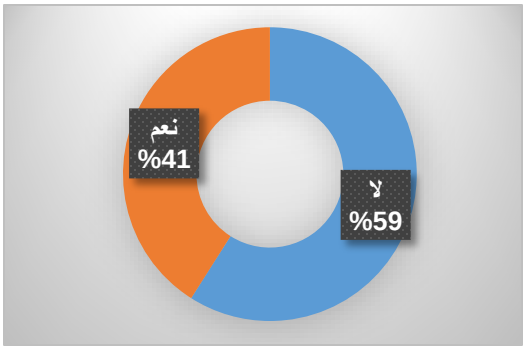
شكل 1. نسبة المشاركين في الاستبيان



1.4 هل قامت مؤسستكم بإعداد خطة أو استراتيجية لاستخدام البيانات الكبيرة؟

عملت حوالي 44% فقط من الجهات المعنية العربية على إعداد خطة استراتيجية لاستخدام البيانات الكبيرة، مما يدعوا للمزيد من الجهد لتعزيز تطبيقات البيانات الكبيرة.

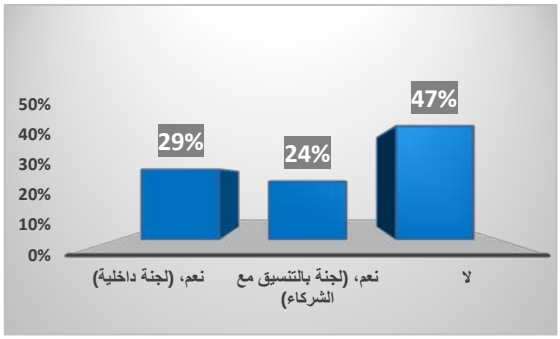
شكل 2. هل قامت مؤسستكم بإعداد خطة أو استراتيجية لاستخدام البيانات الكبيرة؟



2.4 هل تم تشكيل لجنة أو وحدة فنية متخصصة بموضوع البيانات الكبيرة في مؤسستكم؟

قامت حوالي 29% من الجهات المعنية العربية بتشكيل لجنة أو وحدة فنية داخلية متخصصة في موضوع البيانات الكبيرة، في حين 47% ليس لديها لجنة داخلية مختصة، ولا تعمل بالتنسيق مع الشركاء في إطار لجان متخصصة.

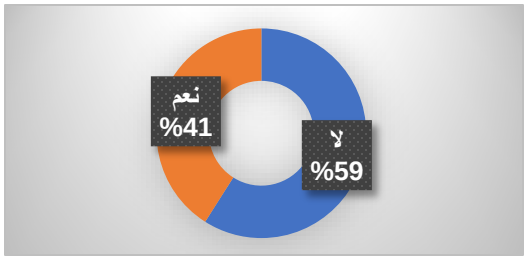
شكل 3. هل تم تشكيل لجنة أو وحدة فنية متخصصة بموضوع البيانات الكبيرة في مؤسستكم؟



3.4 هل لدى مؤسستكم عضوية في لجان إقليمية أو دولية في مجال البيانات الكبيرة

حوالي 41% من الجهات المعنية العربية لديها عضوية في لجان إقليمية أو دولية في مجال البيانات الكبيرة.

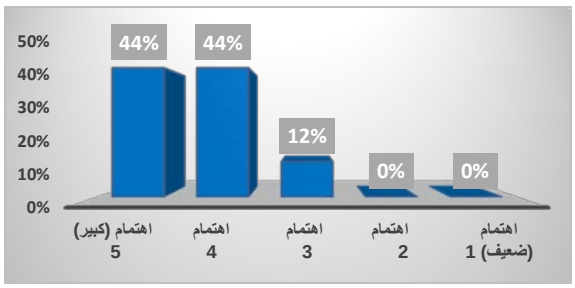
شكل 4. هل لدى مؤسستكم عضوية في لجان إقليمية أو دولية في مجال البيانات الكبيرة؟



4.4 ما هو مستوى إهتمام الإدارة العليا في مؤسستكم بموضوع البيانات الكبيرة؟

تولي الإدارة العليا في أغلبية الجهات العربية المعنية (88%) إهتماماً بموضوع البيانات الكبيرة، ويولي 44% منها إهتماماً كبيراً بالموضوع.

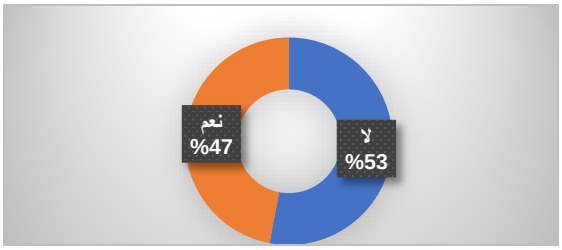
شكل 5. مستوى إهتمام الإدارة العليا في مؤسستكم بموضوع البيانات الكبيرة



5.4 هل تعملون حالياً على أي مشروع للبيانات الكبيرة؟

يعمل حوالي 47% من المؤسسات العربية المعنية على مشاريع لها صلة بالبيانات الكبيرة.

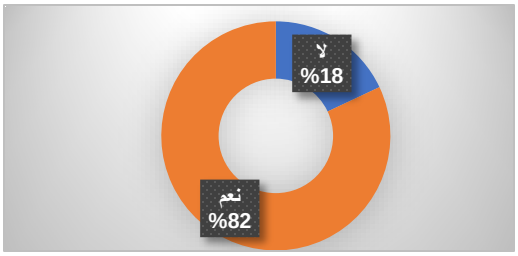
شكل 6. هل تعملون حالياً على أي مشروع للبيانات الكبيرة؟



6.4 هل تنظر مؤسستكم في البدء بمشاريع لها صلة بالبيانات الكبيرة في الفترة المقبلة؟

يسعى حوالي 82% من المؤسسات العربية لبدء العمل على مشاريع لها صلة بالبيانات الكبيرة.

شكل 7. هل تنظر مؤسستكم في البدء بمشاريع لها صلة بالبيانات الكبيرة في الفترة المقبلة؟



7.4 ما الغرض من مشاريع البيانات الكبيرة التي تسعى مؤسستكم لبدئها؟

تحديث عملية الإنتاج الاحصائي والتحليل الاقتصادي وتحسين وتنويع أشكال المخرجات:

- التعاون مع الشركاء الاستراتيجيين في القطاع الخاص والحكومي لاستخدام مصادر البيانات الكبيرة في إنتاج الإحصاءات الرسمية وفقاً لأفضل الممارسات العالمية وتوصيات المنظمات الدولية.
- استخدام البيانات الكبيرة لبناء مؤشرات إحصائية رسمية، واستخراج إحصائيات مكملة من خلال البيانات الكبيرة، وإثراء التحاليل الاقتصادية والمالية، وتوفير منتجات وخدمات جديدة تساعد صناع القرار في الدولة في رسم الخطط المستقبلية.
- دراسة الاستفادة من البيانات غير المهيكلة لتوفير إحصاءات بسرعة عالية وتحسين جودة الانتاج الاحصائي
- دراسة مدى فعالية نظم تحليل البيانات الكبيرة.
- توفير لوحة قيادة لتسريع وتحسين عملية أخذ القرار.
- توفير الأرقام القياسية (أسعار المستهلك CPI، أسعار المنتج PPI، الصادرات والواردات MXPI).

تجميع البيانات وتطوير قواعد البيانات:

- تجميع أكبر عدد من البيانات الإقتصادية خاصة المتعلقة بالنظام المصرفي والمعاملات الخارجية بغرض التحليل والمتابعة وكذلك وضع حجم كاف من المعلومات تحت تصرف الباحثين والمهتمين بالعمل الاحصائي.
- تلبية متطلبات توفير مؤشرات جديدة مثل مؤشرات الأهداف الإنمائية المستدامة (Sustainable Development Goals).
- تطوير قواعد البيانات الإحصائية، والإنتقال إلى قواعد بيانات أكثر تطوراً وتخصصاً للتعامل مع السلاسل الزمنية.

تخفيض التكلفة: تخفيض التكلفة والجهد والوقت في إجراء المسوح وتخفيف العبء عن المستجوبين.

الوصول إلى البيانات: تجميع وتنظيم كافة البيانات في مكان واحد، مما يؤدي إلى سرعة الوصول إلى البيانات وتسهيل عملية إعداد التقارير والدراسات المتعلقة بالدوائر في المؤسسة. وكذلك كشف اتجاه المتغيرات الإقتصادية ونقاط التحول فيها خلال الفترات المختلفة والتنبؤ بها قبل حدوثها، إضافة الى تقييم تأثير القرارات والمخاطر المرافقة لها على النظام الإقتصادي والمالي، وبالتالي تحقيق هدف البنك المركزي في تغطية الاستقرار المالي والرقابة على الجهاز المصرفي، إضافة إلى السياسة النقدية.

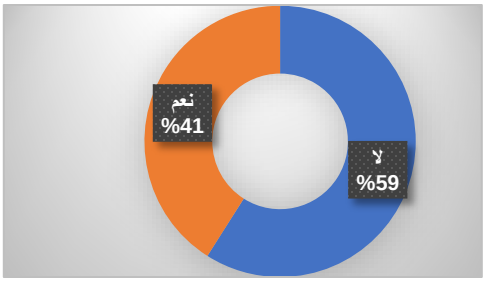
الإنتقال من مشروعات البيانات الكبيرة من مرحلة التجربة إلى مرحلة الإنتاج: دعم مبادرات مشروعات البيانات الضخمة والانتقال بهذه المشروعات من مرحلة التجربة إلى الإنتاج، مما سيساعد على إنتاج إحصاءات أسرع، وتقليل عبء الاستجابة، وتحديث المنتجات من خلال آلية تبسيط وتسهيل عملية جمع الأسعار، والتنوع في أشكال المخرجات، وقياس إنتاج المحاصيل الزراعية باستخدام صور الأقمار الصناعية والتنبؤ بإنتاجها، واستكشاف فائدة البيانات الكبيرة من صور الأقمار الصناعية للأراضي لتعديل وتحسين المعلومات الإحصائية، وتحسين الرعاية الصحية للمرضى، وتتبع أنماط التنقل للمصابين بأمراض وبائية للسيطرة والتحكم في المرض، وقياس كثافة المرور على الطرق السريعة، ومتابعة حوادث السير وتحديد أماكنها، وقياس عدد الأفراد العاملين أو الدارسين في موقع مختلف عن مكان الإقامة.

تطبيق سياسة الحوكمة: المشاركة في تطبيق سياسة الحكومة المتعلقة بمجتمع المعلومات.

8.4 هل تستخدم مؤسستكم مصادر البيانات الكبيرة؟

تستخدم حوالي 41% فقط من المؤسسات العربية المعنية مصادر البيانات الكبيرة.

شكل 8. هل تستخدم مؤسستكم مصادر البيانات الكبيرة؟



ما الدوافع حول عدم استخدام أي مصادر للبيانات الكبيرة؟

توفر البيانات والوصول إليها:

- عدم توفر البيانات غير المهيكلة.
- صعوبة النفاذ إلى مصادر البيانات الكبيرة.
- جاري العمل للاستفادة من مصادر البيانات الكبيرة.

التقنيات: عدم توفر الإمكانات التقنية التي تساعد على استخدام البيانات الكبيرة.

المهارات: قلة الموارد البشرية، والحاجة إلى كفاءات وخبرات كافية في مجال البيانات الكبيرة، وبناء القدرات للتعامل مع هذه البيانات.

التمويل: عدم توفر الإمكانات المادية لتمويل مشاريع البيانات الكبيرة

الإتفاقيات: عدم وجود إتفاقيات تبادل بيانات مع الأطراف المتداخلة.

خصوصية وسرية البيانات الكبيرة: الحاجة إلى وضع لوائح منظمة واعتمادها للتعامل مع خصوصية وسرية البيانات الكبيرة، وكذلك تعميم التوجيهات والإرشادات ذات الصلة بسرية وخصوصية هذه البيانات على مختلف مكونات النظام الإحصائي الرسمية.

المرحلة التجريبية للمشاريع: لم يصل استخدام البيانات الكبيرة مرحلة التنفيذ، والمشاريع في طور دراسة الفرص، وفي المرحلة التجريبية لاستخدام هذه البيانات.

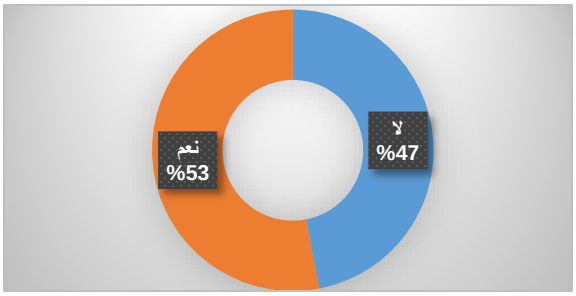
9.4 ما مصادر البيانات الكبيرة التي تستخدمها مؤسستكم؟

شكل 9. مصادر البيانات الكبيرة التي تستخدمها مؤسستكم.

10.4 هل لدى مؤسستكم إمكانية الوصول المباشر إلى مصادر البيانات الكبيرة؟

تتوفر حوالي 53% من المؤسسات العربية المعنية على إمكانية الوصول المباشر إلى مصادر البيانات الكبيرة.

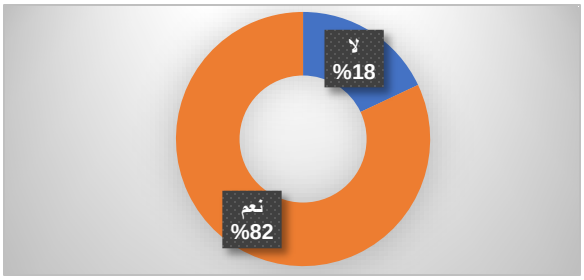
شكل 10. هل لدى مؤسستكم إمكانية الوصول المباشر إلى مصادر البيانات الكبيرة؟



11.4 أ. هل تستخدمون البيانات الإدارية الرسمية؟

حوالي 82% من المؤسسات العربية المعنية تستخدم البيانات الإدارية الرسمية.

شكل 11. هل تستخدمون البيانات الإدارية الرسمية؟

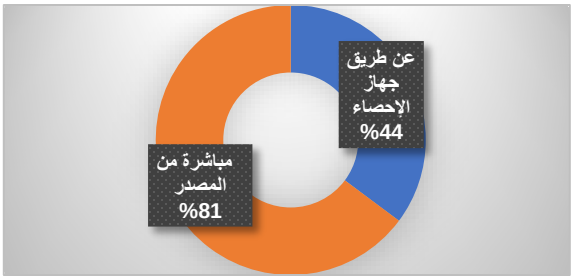


ب. ما هي مصادر البيانات الادارية المعتمدة لديكم؟

شكل 12. مصادر البيانات الادارية المعتمدة

ت. كيف تحصل مؤسستكم على البيانات الإدارية؟

شكل 13. كيف تحصل مؤسستكم على البيانات الإدارية؟



حوالي 81% من المؤسسات العربية المعنية تستخدم البيانات الإدارية الرسمية مباشرة من المصدر.

12.4 ما هي أكبر خمسة تحديات تواجه مؤسستكم في استخدام البيانات الكبيرة أو البدء في مشاريع البيانات الكبيرة؟

أشارت الجهات المستطلعة إلى أهم خمسة تحديات تواجهها في استخدام البيانات الكبيرة أو البدء في مشاريع البيانات الكبيرة، كما يوضح الجدول رقم 4:

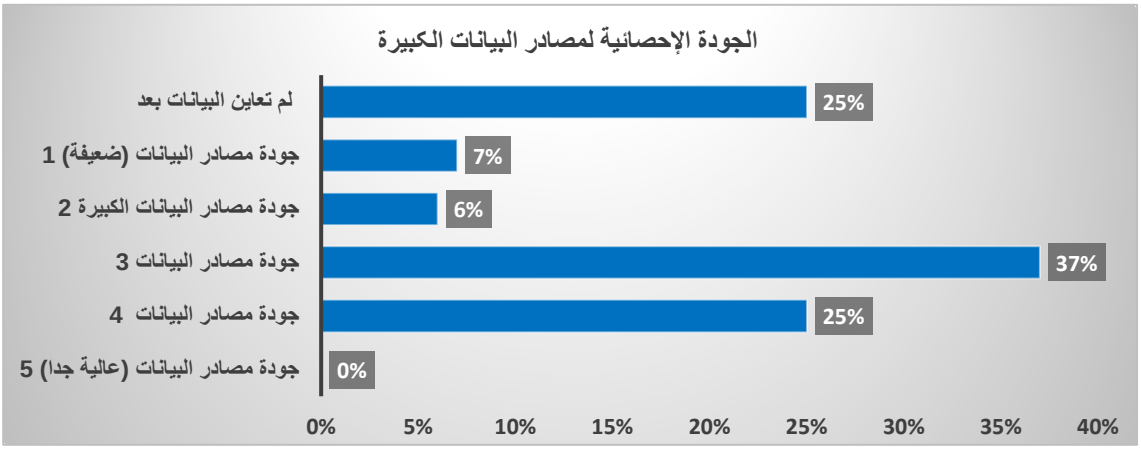
جدول 4: أكبر خمسة تحديات تواجه الجهات العربية المعنية باستخدام البيانات الكبيرة أو البدء في مشاريع البيانات الكبيرة

التحديات	الوصف	النسبة (%)
أكبر تحدي (درجة1)	- إمكانية الوصول إلى البيانات الكبيرة والحصول عليها بصفة دورية والتشريعات المتعلقة بها.	%50
	- المهارات التقنية والخبرات الوطنية في مجال البيانات الكبيرة.	%17
	- التعاون مع الجهات المنتجة للبيانات وإبرام اتفاقيات تبادل البيانات.	%17
	- الإمكانات المادية لبدء مشروعات البيانات الكبيرة وتمويلها.	%16
التحدي (درجة 2)	- صعوبة الحصول على البيانات الكبيرة من مصادرها واستمرارية توفرها، وكذلك حوكمة هذه البيانات ونوعيتها ومدى ملاءمتها وموثوقيتها وربط كافة البيانات في قاعدة البيانات.	%45
	- استراتيجية بناء القدرات والتدريب، والمهارات والكفاءات والتمكين العلمي في مجال البيانات الكبيرة.	%36
	- حماية خصوصية وسرية البيانات.	%10
	- الميزانية اللازمة لهذه المشاريع.	%9
التحدي (درجة 3)	- القدرات المالية.	%27
	- عدم وجود آليات ومنهجيات مستخدمة وسياسات واضحة لاستخدام هذه البيانات.	%27
	- سرية وأمن وسلامة البيانات.	%18
	- الوصول إلى البيانات المسجلة الملكية وتباين طرق تصنيف البيانات سواء المهيكلة أو غير المهيكلة.	%18
	- تكوين فريق مختص في مجال البيانات الكبيرة.	%10
التحدي (درجة 4)	- بناء القدرات التقنية والحاجة إلى الكوادر المتخصصة المهنية.	%25
	- ميدان جديد ومحدودية التجارب العالمية في هذا المجال وضعف التوثيق.	%25
	- سرية البيانات.	
	- البنية التحتية والإدارات التقنية.	%12.5
	- تعزيز الشراكات ذات المنفعة المتبادلة مع أصحاب البيانات (شركاء العمل الاحصائي).	%12.5
	- تحديد حاجيات المؤسسة في مجال البيانات الكبيرة.	%12.5
أقل تحدي (درجة 5)	- اختلاف التصانيف والمعايير المطبقة في الجهات المنتجة للبيانات، وعدم توافر المنهجيات العالمية لاستخدام البيانات الكبيرة في إنتاج الإحصاءات.	%33
	- بناء القدرات البشرية اللازمة للتعامل مع البيانات الكبيرة.	%17
	- كيفية الوصول إلى المصادر الحديثة والجديدة للبيانات الإحصائية وتطبيقها لخلق نموذج أعمال للشراكة- تحديد أصحاب المصلحة الرئيسين في القطاع الخاص.	%17
	- زيادة استخدام الاستبيانات وتقنيات التنبؤ.	%17
	- إدارة التغيير في إطار تطبيقات البيانات الكبيرة.	%16

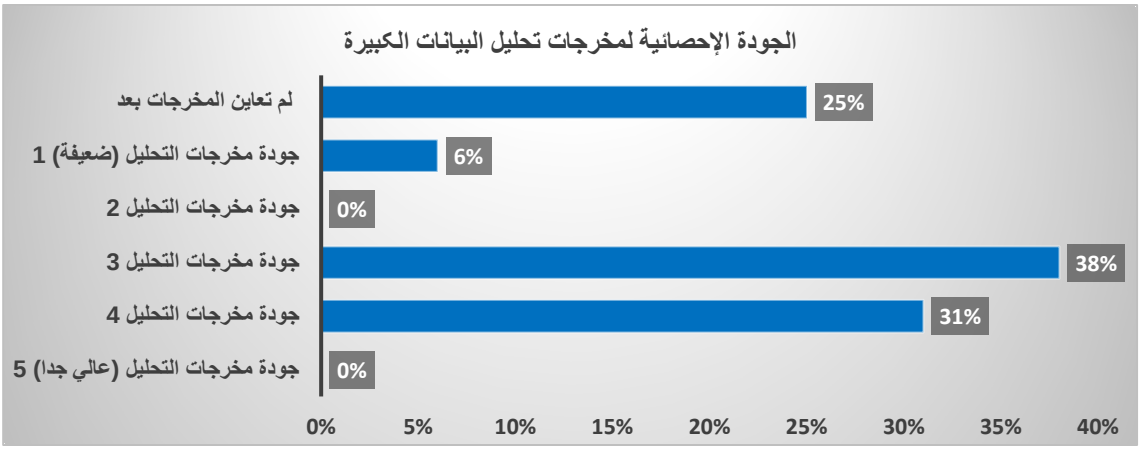
13.4 ما هو تقييمكم للجودة الإحصائية لمصادر ومخرجات تحليل البيانات الكبيرة؟

أشار حوالي 25% من الجهات المستطلعة إلى عدم معاينتها للبيانات الكبيرة ولمخرجاتها، بينما أوضحت الجهات الأخرى تقييمها الإيجابي لجودة البيانات الكبيرة ومخرجات تحليلها.

شكل 14. ما هو تقييمكم للجودة الإحصائية لمصادر البيانات الكبيرة؟



شكل 15. ما هو تقييمكم للجودة الإحصائية لمخرجات تحليل البيانات الكبيرة؟



14.4 ما المخرجات التي تتوقعونها نتيجة استخدام البيانات الكبيرة (على سبيل المثال القدرة على التنبؤ الآني) - إن توفرت؟

توفير إحصاءات ومنتجات وخدمات جديدة

- إنتاج مؤشرات جديدة وتلبية متطلبات توفير مؤشرات جديدة مثل مؤشرات الأهداف الإنمائية المستدامة.
- توفير إحصائيات شبه آنية، وتوفير معلومات آنية تتواءم مع الطلب المتزايد من متخذي القرار .
- إنتاج مؤشرات إحصائية على مستويات جغرافية صغيرة، وإنتاج مؤشرات جديدة.
- تطوير مؤشرات تقنيات المعلومات والإتصالات وإعداد مؤشرات جديدة واستخدامات جديدة في مجال الاحصاء السكاني من خلال بيانات الهاتف المحمول.
- العديد من المجالات الحيوية ذات الطبيعة الديناميكية المرتفعة تتوفر بياناتها من خلال البيانات الكبيرة التي توفر الفرصة لبناء قدرات تنبؤية مثل معلومات سوق العمل.
- استخدام تقنيات إحصائية متطورة في تحليل البيانات.
- إتاحة سلاسل زمنية من البيانات أطول بغرض التنبؤ والتحليل.
- القدرة على اتخاذ قرارات آنية وبأكثر فاعلية.

تخفيض التكلفة وتحديث عملية الإنتاج الإحصائي

- تخفيض التكلفة على المدى المتوسط (تكلفة المسوح وعملية الانتاج).
- تخفيف العبء عن المستجوبين.
- تحديث عملية الإنتاج الإحصائي.

تحسين جودة ودقة وسرعة الإحصاءات والدراسات

- تحسين جودة النماذج الإحصائية والمؤشرات الإحصائية، ومخرجات المنتجات الاحصائية مما يساهم في صناعة قرارات تنموية أفضل.
- توفير إحصاءات متسقة وبسرعة عالية وفي الوقت المناسب لأغراض صنع القرار.
- توفير رؤية أكثر وضوحاً ودقة للمحللين.
- إجراء دراسات مالية واقتصادية، تساعد متخذي القرار على رسم السياسات الإدارية أو الإقتصادية المناسبة والقدرة على التنبؤ بالمتغيرات الإقتصادية قبل حدوثها واسترجاع البيانات التجمعية في الوقت المناسب لاتخاذ القرارات المناسبة.
- التنبؤ الآني والتحليل التنبئي.
- التأثير على نتائج البحوث.
- المساعدة على الفهم الصحيح للأحداث الآنية وتعزيز الوعي بالأحداث.
- البيانات الكبيرة ميزة تنافسية للشركات، وهو ما يظهر في تقييم قيمة الشركات وأصولها من البيانات وقدراتها على التحليل.

15.4 ما القضايا المرتبطة بالسرية والخصوصية التي تواجهها مؤسستكم فيما يتعلق بمصادر البيانات الكبيرة؟

عدم وضوح الرؤية في التعامل مع هوية البيانات وجانب السرية والخصوصية المرتبطة بها.

قضايا السرية والخصوصية في تحليلات البيانات الكبيرة من المخاوف الرئيسية على مستوى العالم

- الحاجة إلى تطوير مجموعة من المبادئ التوجيهية أو البروتوكولات التي من شأنها تعزيز إمكانية الوصول إلى بيانات القطاع الخاص واستخدامها، نظراً للأسباب التالية:
- تحفظ وتخوف منتجي البيانات فيما يخص المعلومات الشخصية للعملاء مما يشكل خرق للثقة بينهم وبين العميل، حيث يعتبر تدخلاً في الحقوق الفردية، والخصوصية، والهوية.
- حساسية وسرية بعض المعلومات تمنع المصادر المعنية من الإفصاح عنها، مما قد يؤثر على نتائج تحليل البيانات.
- صعوبة الحصول على بيانات القطاع الخاص في ظل غياب تشريعات حوكمة استخدام البيانات الكبيرة على المستوى المحلي وارتباطها بتشريعات دولية للشركات العالمية.
- صعوبة تبادل المعلومات نظراً لقواعد السرية والخصوصية للمستخدمين.

التعقيدات التقنية والتحديات القانونية

- التعقيدات التقنية التي تصاحب تبادل المعلومات والتي تضمن سرية البيانات خاصة على مستوى الجهات غير الحكومية.
- تحديد التحديات القانونية التي لا يغطيها قانون الاحصاءات العامة.
- الجوانب العملية غير المعروفة حالياً بسبب عدم التطبيق أو التطبيقات متواضعة للبيانات الكبيرة.

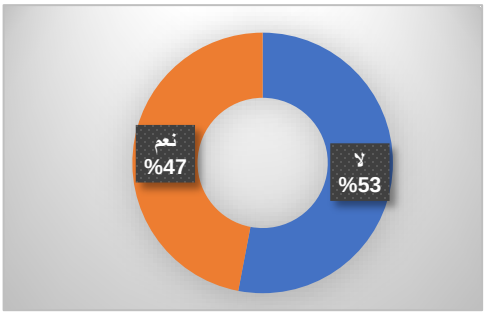
إقامة شراكات مع القطاع الخاص

- الحاجة إلى إقامة شراكات مع القطاع الخاص يُلبي الشواغل المتعلقة بالخصوصية واحترام قواعد السرية.
- ضرورة الحفاظ على سرية المعطيات الفردية لمؤسسات الإقراض والمؤسسات المالية الأخرى وبعض المؤسسات الغير المالية. والبيانات السرية لدى الجهات التي تشرف عليها الجهات المسؤولة والتي تتسم بالسرية العالية مثل البيانات البنكية للعملاء.
- الحفاظ على سرية وخصوصية بيانات الأفراد مثل الاسم ورقم الهاتف والبيانات المعرفة للأفراد.

16.4 هل توفر لمؤسستكم إطار قانوني لمعالجة قضايا السرية والخصوصية في البيانات الكبيرة؟

يحتاج حوالي 53% من المؤسسات المستطلعة إلى إطار قانوني لمعالجة قضايا السرية والخصوصية في البيانات الكبيرة.

شكل 16. هل لديكم إطار قانوني لمعالجة قضايا السرية والخصوصية في البيانات الكبيرة؟



17.4 ما المهارات التي تعتبرونها الأهم بالنسبة للموظفين الذين يتعاملون مع البيانات الكبيرة؟

مهارات التحليل والابتكار وعلم البيانات

- إتقان أدوات وطرق التنبؤ.
- القدرة على الإبداع والابتكار، والقدرة على التحليل المتعمق للبيانات، والقدرة على طرح الحلول السريعة، والقدرة على التركيز في فهم العلاقات والروابط.
- القدرة على استخدام المخرجات التي توفرها نظم البيانات الكبيرة (مثل تحليل نتائج النماذج الإحصائية).
- مهارات في علوم الاحصائيات، وعلوم البيانات وعلوم وهندسة المعلوميات.
- القدرة على تنقيب البيانات (Datamining).
- "Data science".
- "Machine learning".
- "Systems Analysis".
- القدرة على تحديد الانحياز في البيانات واكتشاف الأخطاء.
- تطوير الجانب الأكاديمي للمستخدمين.

البرمجة واستخدام أنظمة/آليات تقنية المعلومات

- القدرة العالية على استعمالات تقنيات ونظم المعلومات والبرمجة، وإتقان استخدام أدوات البيانات الكبيرة.
- المعرفة بالتقنيات المتعلقة بالبيانات الكبيرة مثل (Hadoop, Spark، وغيرها).

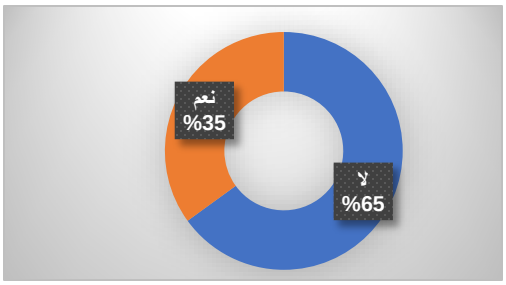
تجميع ونقل وتخزين ومعالجة البيانات

- الإلمام بالأدوات المستخدمة في معالجة البيانات الكبيرة.
- معرفة آليات وبرمجيات نقل البيانات الكبيرة.
- المعرفة بأنظمة إدارة وتشغيل وتخزين البيانات الكبيرة.
- القدرة على تجميع البيانات.
- قواعد البيانات (NoSQL).
- إدارة تدفق البيانات.

18.4 هل تقومون بتوظيف خبراء متخصصين في مجال البيانات الكبيرة؟

حوالي 35% فقط من الجهات المعنية تقوم بتوظيف خبراء مختصين وتدريب موظفيها في مجال البيانات الكبيرة.

شكل 17. هل تقومون بتوظيف خبراء متخصصين في مجال البيانات الكبيرة؟



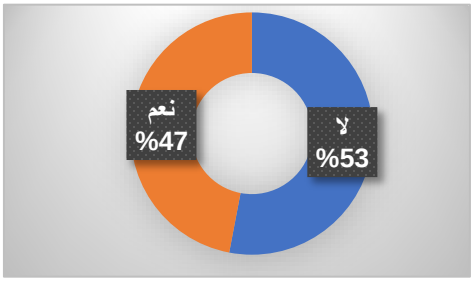
19.4 كيف عالجتم و/أو تخططون لمعالجة النقص في المهارات في البيانات الكبيرة لدى مؤسستكم؟

- تدريب وتأهيل الموارد البشرية، ووضع مسار وظيفي ملائم ودعم خبرات الموظفين من خلال شراكات مع الشركات العالمية والإطلاع على أفضل الممارسات، وتبادل الخبرات الدولية.
- وضع خطة عمل للتواصل مع مجتمع شباب علماء البيانات ودعمهم.
- العمل على استقطاب الشباب في مجالات تقنية المعلومات وعلوم الحاسب وتشجيعهم على الاندماج في مجال الاحصاء وعلم البيانات.
- استقطاب مزيد من المختصين في هذه المجالات، والاستعانة بخبراء متخصصين في البيانات.
- المشاركة في الفعاليات الدولية المتخصصة من ندوات وملتقيات.
- تنظيم ورش عمل لتطوير المهارات في استخدام البيانات الكبيرة.
- المشاركة في فرق العمل الدولية والتجارية.
- عدم وجود إمكانيات لمعالجة النقص في المهارات في البيانات الكبيرة حالياً.

20.4 هل لديكم اتفاقيات تعاون و/أو إجراءات عمل لشراكات مع مزودي البيانات الكبيرة؟

حوالي 47% من المؤسسات العربية المعنية لديها اتفاقيات تعاون/ إجراءات عمل لشراكات مع مزودي البيانات الكبيرة

شكل 18. هل لديكم اتفاقيات تعاون و/أو إجراءات عمل لشراكات مع مزودي البيانات الكبيرة؟



21.4 هل مؤسستكم على استعداد للمشاركة مع دول او مؤسسات أخرى أو لجان أو فرق عمل في مجال البيانات الكبيرة؟

عبرت كل الجهات المعنية المستطلعة عن استعدادها للمشاركة مع دول أو مؤسسات أخرى أو لجان أو فرق عمل في مجال البيانات الكبيرة.

أ. ما مجالات مشاركتكم مع الجهات الأخرى؟

مشاريع ذات الصلة بالبيانات الكبيرة

- المشاركة في مشروع تجربة استخدام البيانات الكبيرة في إنتاج الاحصاءات الرسمية (مركز الاحصاء لدول مجلس التعاون لدول الخليج العربية وفريق عمل البيانات الكبيرة).
- مشروع جمع الأسعار – استخدام صور الأقمار الصناعية – استخدام البيانات المكانية للهواتف المحمولة.
- تطوير مؤشرات تقنية المعلومات والإتصالات (الاتحاد العالمي للإتصالات (ITU)).
- إدارة مشاريع البيانات الكبيرة - إدارة تدفق البيانات - التحليل - عرض ومشاركة البيانات في إطار الاستفادة من كافة الأطراف المهتمة في انتظار تجسيد هذه الرؤية داخلياً.

بناء الخبراء وتطوير طرق التحليل

- بناء الخبرات والمهارات في مجال استخدام البيانات الكبيرة في الاحصاء (الفريق العالمي للبيانات الكبيرة - الأمم المتحدة).
- مجال التحليل واستخدام التقنيات الحديثة.

عضوية في لجان وفرق عمل، والمشاركة في ورش عمل

- لجنة المنصة العالمية للبيانات والخدمات والتطبيقات (الفريق العالمي للبيانات الكبيرة - الأمم المتحدة).
- لجنة "IRVING FISHER" التابعة لبنك التسويات الدولية، حيث تهتم اللجنة بمواضيع الإحصاءات للبنوك المركزية بما في ذلك البيانات الكبيرة.
- عضو في فرق البيانات الكبيرة التابع للمركز الإحصائي لدول مجلس التعاون لدول الخليج العربية.
- المشاركة في ورش عمل .

الاستفادة من أفضل الممارسات وتبادل الخبرات والبيانات المتوفرة

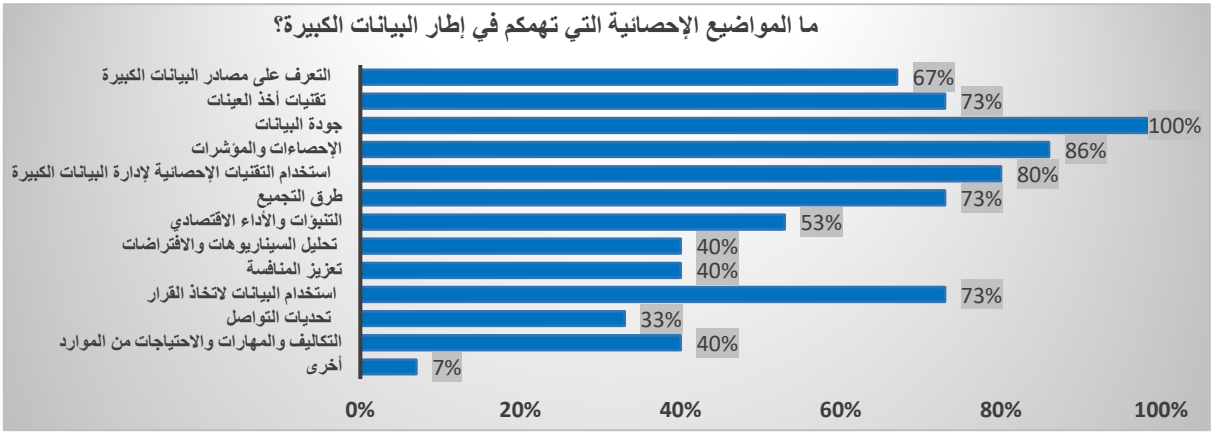
- رصد التطور التقني في مجال البيانات الكبيرة وتبادل الخبرات والتجارب، والاستفادة من التجارب الإقليمية والدولية في جميع المواضيع ذات الصلة بالبيانات الكبيرة.
- الاستفادة من البيانات المتوفرة لدى المؤسسات.
- عرض البيانات الحالية والتفكير في طرق استغلالها بالاستفادة من تجارب بلدان أخرى.

22.4 هل مؤسستكم على استعداد للمساهمة في خارطة طريق و/أو خطة للتعاون الإقليمي حول استخدام وتطوير البيانات الكبيرة؟

عبرت جل الجهات المعنية عن استعدادها للمساهمة في خارطة طريق/أو خطة للتعاون الإقليمي حول استخدام وتطوير البيانات الكبيرة.

23.4 ما المواضيع الإحصائية التي تهم الجهات المعنية في إطار البيانات الكبيرة؟

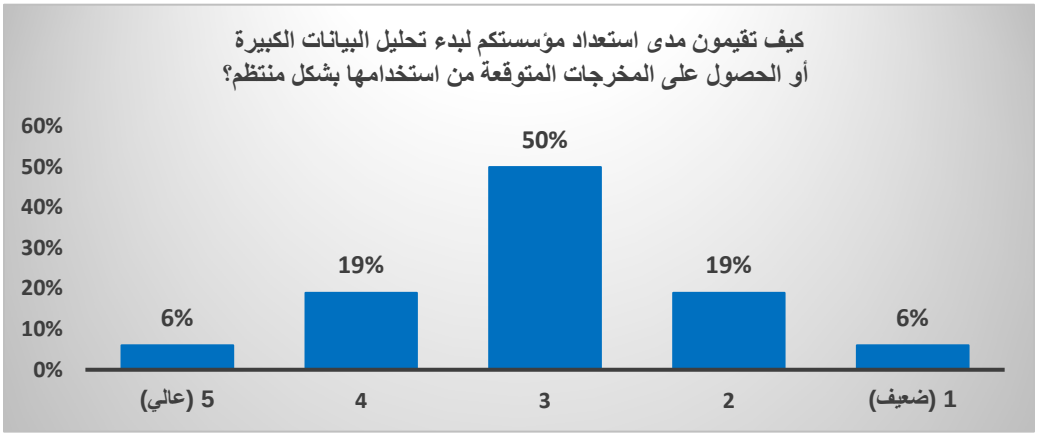
شكل 19. المواضيع الإحصائية التي تهم الجهات المعنية في إطار البيانات الكبيرة



تتصدر جودة البيانات تركيز واهتمام الجهات المستطلعة (100%)، تليها الإحصاءات والمؤشرات (86%)، في حين حوالي 67% من الجهات يتطلع للتعرف على مصادر البيانات الكبيرة.

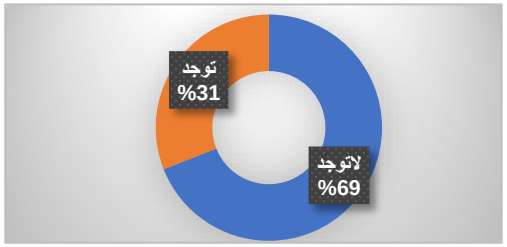
24.4 كيف تقيمون مدى استعداد مؤسستكم لبدء تحليل البيانات الكبيرة أو الحصول على المخرجات المتوقعة من استخدامها بشكل منتظم؟

شكل 20. تقييم مدى استعداد مؤسستكم لبدء تحليل البيانات الكبيرة أو الحصول على المخرجات المتوقعة من استخدامها بشكل منتظم؟



25.4 ما الأدوات التي تستخدمونها لتحليل البيانات الكبيرة؟

شكل 21. التوفر على أدوات لتحليل البيانات الكبيرة



حوالي 31% فقط من المؤسسات العربية المعنية لديها أدوات لاستخدامها في تحليل البيانات الكبيرة.

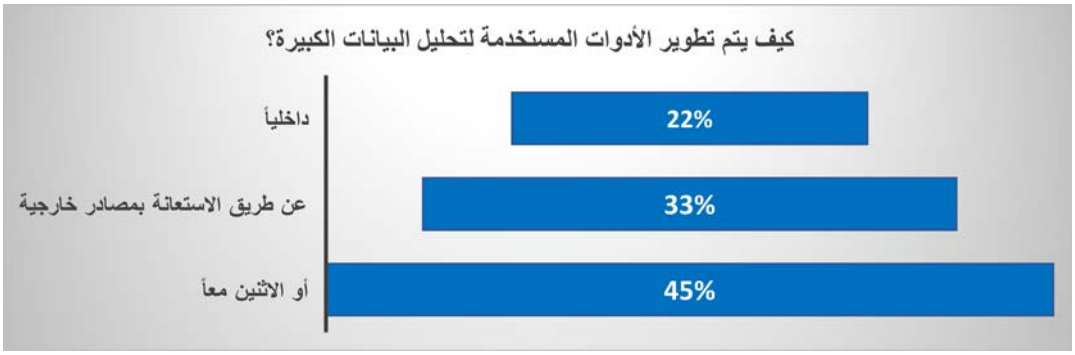
جدول 5: الأدوات التي تستخدمها الجهات المستطلعة لتحليل البيانات الكبيرة

ما الأدوات التي تستخدمونها لتحليل البيانات الكبيرة؟							
"R Language" "R-Studio"	"Alteryx" "QlickView" "KNIME" "Cloudera"	نظام "Fame" المتخصص في السلاسل الزمنية	Microstrategy	Hadoop and Postgre SQL and ShellScripting	التطلع لاستخدام "Tab lean " كأداة للتحليل وعرض البيانات	بعض التطبيقات الإحصائية المتعلقة باستخراج وإعداد البيانات من "data "Warehouse	البرامج الإحصائية (Eviews, SPSS...)

26.4 كيف يتم تطوير الأدوات المستخدمة لتحليل البيانات الكبيرة؟

يعمل حوالي 22% فقط من الجهات المستطلعة على تطوير الأدوات المستخدمة لتحليل البيانات الكبيرة داخلياً، بينما يلجأ 33% منها للاستعانة بمصادر خارجية.

شكل 22. كيف يتم تطوير الأدوات المستخدمة لتحليل البيانات الكبيرة؟



27.4 يرجى وصف بعض الآثار المحتملة للموارد على مؤسستكم في بدء تحليل البيانات الكبيرة و / أو الحصول على المخرجات المتوقعة من استخدامها بشكل منتظم، لا سيما من حيث موارد تقنية المعلومات (كل من البرمجيات والأجهزة) ورأس المال البشري (فيما يخص تقنية المعلومات، والإحصاءات، والبحوث الاقتصادية، وعمليات التصنيف، وأمن البيانات، والأمن الوطني).

أشارت الجهات المستطلعة إلى عدد من الآثار المحتملة الناتجة عن تطبيقها للبيانات الكبيرة وعمليات تحليل هذه البيانات، والتي تشمل مجالات مختلفة على رأسها قضايا البنية التحتية التقنية، والموارد البشرية الضرورية، والبرمجيات والأدوات اللازمة، والموارد المالية الكافية، كذلك معالجة قضايا التشريعات والقوانين.

في نفس الوقت، وصفت الجهات المستطلعة البيانات الكبيرة كأداة فعالة لدعم الإحصاءات الرسمية وتحليلات الإقتصاد الكلي، لاسيما لتنويع المخرجات وتوفير إحصاءات ومؤشرات جديدة، كذلك لتحسين جودة ودقة الإحصاءات وسرعة الحصول عليها، إضافة إلى تحديث عمليات الإنتاج والعمل على دراسات متخصصة لاتخاذ القرارات المناسبة.

البنية التحتية التقنية

- يتطلب الوصول الى البيانات الكبيرة الاستثمار في البنية التحتية التقنية المتعلقة باستخراج البيانات ومعالجتها وكذلك في المصادر البشرية مما سيرفع من المستوى:

الموارد البشرية:

- الحاجة إلى مختصين وخبراء في علم البيانات وتحليلها.
- الحاجة إلى تطوير الموارد البشرية في مجالات متخصصة ومتعددة.
- نقص في الخبرات الفنية والعملية للموارد البشرية.
- لا يوجد في الهيكلية الحالية قسم متخصص معني بالبيانات الكبيرة.

البرمجيات والأدوات اللازمة:

- الحاجة إلى البرامج والأدوات اللازمة لحفظ وتخزين وتحليل البيانات ودمج كافة بيانات المؤسسة في مكان مركزي.
- البرمجيات بالأجهزة تحتاج لتطوير وترقية مستمرة نظراً للزيادة المتوقعة في حجم البيانات.
- الحاجة إلى توظيف تقنيات مختلفة للحفاظ على أمن المعلومات.

التمويل

- استخدام البيانات الكبيرة مجدي على المدى المتوسط والمدى البعيد من حيث تخفيض التكلفة وتوفير البيانات خلال فترات زمنية قصيرة (تحسين الوقتية)، إلا أن هناك:
- حاجة إلى موارد تساعد على تطوير التقنيات المعلوماتية الضرورية لتطوير أنظمة البيانات الكبيرة، والوصول إلى هذه البيانات وتحليلها، حيث أن الموارد تشكل عائقاً في هذا المجال.

التشريعات والقوانين: الحاجة إلى تطبيق قوانين لحفظ سرية وخصوصية المعلومات الحساسة.

التعاون والشراكات: الحاجة إلى التعاون مع المؤسسات الخارجية في مجال البيانات الكبيرة، والتعاقد مع شركاء متعددين لتوفير خدمات مختلفة مثل تخزين البيانات.

تحديث عمليات الإنتاج وتنويع المخرجات

- استخدام البيانات الكبيرة يقوم أيضاً على تسريع عملية إنتاج الإحصاءات وزيادة جودتها ودقتها.
- الحصول على مخرجات يمكن من خلالها تحليل الوضع الحالي والقدرة على التنبؤ وتحسين أساليب القياس، وزيادة مهارات العاملين على قواعد البيانات، والقدرة على توفير الإحصاءات اللازمة لإعداد الدراسات والبحوث الإقتصادية المرتبطة بمجالات العمل.
- القيام بدراسات متخصصة.
- إضافة مجالات جديدة لكيفية الاستفادة من البيانات المتوفرة، على سبيل المثال من واقع استخدام الهواتف النقالة.

28.4 هل لديكم مقترحات أخرى لتطوير تطبيقات البيانات الكبيرة، من تشريعات، وموارد بشرية، وتقنيات المعلومات، والوصول إلى البيانات؟

شملت مقترحات الجهات المستطلعة في الدول الأعضاء العديد من المحاور، تم تجميعها في الشكل 22.

- إعداد برنامج متكامل لمساعدة المؤسسات العربية على اكتساب تقنيات البيانات الكبيرة الواعدة، من خلال:

- تنظيم دورات تكوينية منتظمة في مجال البيانات الكبيرة لتأهيل وتدريب متخصصين في البيانات الكبيرة، والعمل على تمكين جيل جديد من خبراء علم البيانات.
- بعثات مساندة فنية ومهمات لتبادل الخبرات والكفاءات.
- مساعدة ومرافقة المؤسسات المبتدئة في هذا المجال.
- نقل التجارب المعرفية للأنظمة التطبيقية والتحليلية والتقنية.
- اختيار التطبيقات والأدوات المناسبة والمطلوبة.
- تقديم الدعم من قبل صندوق النقد العربي ومؤسسات دولية لمساعدة الجهات العربية على البدء في مشاريع البيانات الكبيرة وتحفيز الدول الأعضاء وأجهزة الإحصاء والمصارف المركزية على تطوير استراتيجياتها في القيام بمشاريع ذات الصلة بالبيانات الكبيرة. فالجهات المعنية، تتطلع إلى الاستفادة من البيانات الكبيرة سواءً في الوصول إلى هذه البيانات والمعلومات المطلوبة بمختلف الطرق أو في إمكانية تقليل تكاليف مشاريع البيانات الكبيرة.
- الاستفادة من الجامعات والمراكز البحثية لتطوير برامج تدريبية للتعامل مع البيانات الكبيرة.

- المساعدة على تطوير البنية التحتية القانونية والتقنية.
- إيجاد حلقة وصل بين المنتجين والمستخدمين والمستفيدين من البيانات الكبيرة من خلال إنشاء مجموعات وفرق عمل على المستوى العربي والدولي.
- توفير السياسات المناسبة لتبادل البيانات الكبيرة بالأخص مع القطاع الخاص، حيث ينتج معظم البيانات الكبيرة.
- نشر الوعي حول أهمية البيانات الكبيرة من خلال مؤتمرات وفعاليات وورش عمل عربية. كذلك زيادة وعي بالأجهزة الإحصائية للاستفادة من البيانات الكبيرة في الإحصاءات الرسمية، وتنظيم اجتماعات وندوات ومحاضرات لصناع القرار والقادة، بهدف زيادة حلقة المؤيدين لاستخدام البيانات الكبيرة في الإحصاءات الرسمية.
- تبادل ونقل الخبرات بين الدول العربية للوصول إلى المستوى العالمي.
- إطلاق المشروعات "Pilots" لتطوير المنهجيات والخبرات المطلوبة لتطوير المنتج الاحصائي الرسمي.
- تنفيذ مشاريع البيانات الكبيرة في المؤسسات الإحصائية استناداً إلى دراسة للاحتياجات الفعلية والآفاق التي يفتحها تنفيذ هذه المشاريع والميزات المترتبة على ذلك خصوصاً تحسين المخرجات الإحصائية، وإمكانية الدمج مع بيانات أخرى، والمساهمة في تخفيف العبء عن المستجوبين، إضافة الى تلبية معايير الجودة مثل الدقة والصلة بالواقع وآنية التنبؤ.
- تحديد موقف كل دولة عربية في مجال البيانات الكبيرة، فيما يتعلق بثورة البيانات من أجل التنمية المستدامة، الأمر الذي يتطلب الإستثمار والقيادة والابتكار .
- اتخاذ قرار على مستوى المؤسسات الإحصائية العربية بإدراج موضوع استخدام البيانات الكبيرة ضمن خططها الوطنية نظراً للأهمية العالية التي يحتلها هذا الموضوع حالياً وفي المستقبل.
- مناقشة وتحليل مصادر البيانات الضخمة المتاحة لكل دولة وكيفية استخدامها في الإحصاءات الرسمية وربطها مع مؤشرات التنمية المستدامة.
- العمل على بناء علاقات تعاون وشراكات استراتيجية مع مختلف المؤسسات ذات العلاقة.
- العمل على وضع واعتماد لوائح للتعامل مع خصوصية وسرية البيانات الكبيرة، ومعالجة قضايا التشريعات اللازمة لمشاريع البيانات الكبيرة.

تطبيقات البيانات الكبيرة في الدول العربية

- حصر أفضل الممارسات والاستفادة من التجارب الناجحة في مجال استخدام البيانات الكبيرة، بما فيها الإحصاءات الرسمية.
- المساعدة في توفير موارد مالية مستمرة في إطار تطوير المنظومة المعلوماتية، واستكشاف نظم البيانات الكبيرة بغرض تركيزها بصفة تدريجية.
- إنشاء بنك الأفكار العربي لمشروعات استخدام مصادر البيانات الكبيرة في الإحصاءات الرسمية وربطها بمشروعات التنمية المستدامة لسهولة تبادل المقترحات والأفكار والخبرات بين الدول، ودعم هذا المقترح بأهمية المثلث المكون من ثلاث أجزاء هامة:
  - مصادر البيانات الكبيرة.
  - الإحصاءات الرسمية.
  - مؤشرات التنمية المستدامة.

تم دعم هذا المقترح من قبل الجهات المعنية استناداً لأفضل الممارسات في مجال البيانات الكبيرة في دول أخرى، على سبيل المثال:

- استخدام بيانات التواصل الاجتماعي في دول مثل هولندا، لقياس ثقة المستهلك والمشاعر الأساسية للمواطنين (منفعل، حزين، سعيد، لين، غضبان)، لكن لم يتم ربط ذلك بالإحصاءات الرسمية وبمؤشرات التنمية المستدامة.
- استخدام تتبع أنماط التنقل لمستخدمي "تويتر" (Twitter)، واستخدام بيانات "تويتر" لأغراض إحصاءات السياحة الداخلية في دولة المكسيك.
- كما تم اقتراح وضع الدول لمعايير قبل استخدام مصادر البيانات الكبيرة لأغراض إصدار الإحصاءات الرسمية وربطها بمؤشرات التنمية المستدامة.

شكل 23. مقترحات الجهات المستطلعة في الدول الأعضاء لتطوير تطبيقات البيانات الكبيرة من تشريعات، وموارد بشرية، وتقنية المعلومات، والوصول إلى البيانات وغيرها.

29.4 مشاريع البيانات الكبيرة - الاسم والوصف

أشارت الجهات المستطلعة إلى عدد من مشاريع البيانات الكبيرة الراهنة والمستقبلية.

جدول 6. مشاريع البيانات الكبيرة - الاسم والوصف

مشروع تنفيذ برنامج عمل إحصاءات السياحة (المرحلة الثانية 2016-2017) (مملكة البحرين: هيئة المعلومات والحكومة الالكترونية): إنشاء قاعدة بيانات إحصائية شاملة لإحصاءات السياحة أحد المتطلبات الأساسية لمتخذي القرار وصنّاع السياسات في هذا المجال، وهي عملية تحتاج لضمان واستمرار تدفق البيانات الإحصائية بصورة محدثة وآنية وتكون ذات جودة عالية.

استخدام البيانات الكبيرة كمصدر بيانات جديدة لمؤشرات تقنية المعلومات والاتصالات من خلال طرق مبتكرة (دولة الإمارات العربية المتحدة: الهيئة الاتحادية للتنافسية والإحصاء)،

استخدام نظام "Fame" المتخصص في السلاسل الزمنية بديلاً لقاعدة البيانات الحالية "Mainframe " (دولة الكويت: البنك المركزي الكويتي).

مشروع "Data Warehouse" (المملكة العربية السعودية: مؤسسة النقد العربي السعودي).

منظومة معلومات ذكاء الأعمال -مختبر البيانات الكبيرة (الجمهورية التونسية: البنك المركزي التونسي).

استغلال بيانات مواقع التواصل الاجتماعي لرصد الهدف 16 من أهداف التنمية المستدامة (SDG 16) بشأن السلام والعدالة باستخدام البيانات الكبيرة: تجربة بالتعاون مع برنامج الأمم المتحدة للتنمية (الجمهورية التونسية: المعهد الوطني للإحصاء).

استراتيجية وطنية لبناء نظام احصائي شامل: يضم كل المؤسسات ذات العلاقة وبناء السجلات الإدارية والربط الالكتروني بين المؤسسات (المملكة الأردنية: دائرة الإحصاء العامة).

إنتاج مؤشرات إحصائية من واقع استخدام بيانات الهواتف النقالة (سلطنة عمان: المركز الوطني للإحصاء والمعلومات NCSI).

مشروع أولي للآليات والتطبيقات المناسبة لتطبيق مفهوم البيانات الكبيرة لدعم الخدمات والمنتجات الإحصائية التي تقدمها الهيئة العامة للإحصاء (المملكة العربية السعودية: الهيئة الإحصائية العامة).

جمع بيانات الأرقام القياسية لأسعار المستهلكين: عقد بروتوكول باستلام قائمة بأسعـــار السلع الغذائية والصناعية للمستهلكـين المنتجة لنشرة الأرقام القياسية مع ثلاثة مصادر (استرشادية) عن طريق الانترنت، جاري زيادة المصادر إلى عشرة مصادر من الإطار، وتطوير آليات جمع الأسعار حيث لا تتوقف دقة البيانات على الباحث وتقليل العبء على المصدر وتلافي الأخطاء بعدم تكرار الأسعار أو تثبيتها (جمهورية مصر العربية: الجهاز المركزي للتعبئة العامة والاحصاء).

استخدام صور الأقمار الصناعية في قياس انتاج المحاصيل الزراعية والاستفادة من صور الأقمار الصناعية في الري الحقلي وهو إدارة وترشيد وتوزيع مياه الري الحقلي (جمهورية مصر العربية: الجهاز المركزي للتعبئة العامة والاحصاء).

البيانات الكبيرة والرعاية الصحية:

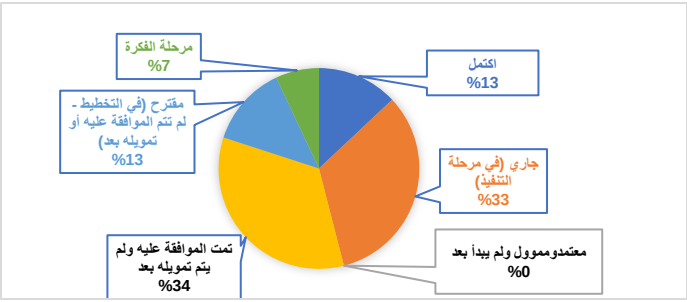
- تتبع أنماط التنقل للمرضى المصابين بأمراض وبائية: مثل الالتهاب الكبدي بواسطة البيانات المكانية للهواتف المحمولة للسيطرة على المرض وعدم انتقال المرض إلى أشخاص آخرين، وتحديد أسباب المرض وتحديد نوع العلاج والرعاية الصحية (جمهورية مصر العربية: الجهاز المركزي للتعبئة العامة والاحصاء).
- تستخدم البيانات الكبيرة في إيجاد سبل لخفض تكاليف الرعاية الصحيّة، وإطلاق مبادرة بعنوان (Cancer Link) عن البيانات الكبيرة التي تُستخدَم في مكافحة مرض السرطان: تقوم فكرة المبادرة على جمع البيانات عن كلّ مريض سرطان في الدولة من أجل تحليل وتحديد التوجّهات أو الأنماط، ومنافع البيانات الكبيرة سوف تأتّى من التركيز الكبير على النتائج التي يتوصّل إليها المريض والإنفاق على المرض، ومعرفة نوع الخدمة المقدمة مثل (علاج التصوير الطبي، والعلاج الكيميائي، والعلاج الطبيعي، والعلاج الإشعاعي، علاج الأورام بالإشعاع والطب النووي). يهدف المشروع إلى الاستفادة من طرق الاصابة للمرضى السابقين لتفادي الاصابة بالمرض والاستفادة أيضاً من إجابات المرضى واستخدامها كمعيارٍ لقياس النتائج الصحية وتقديم الرعاية الصحية (جمهورية مصر العربية: الجهاز المركزي للتعبئة العامة والاحصاء).

البيانات الكبيرة وإحصاءات النقل: حيث أن أحد المجالات الأربعة التي نشرتها مبادرة النبض العالمي للأمم المتحدة هي تنقل البشر ووسائل النقل، استخدام البيانات الكبيرة (البيانات المكانية لمواقع الهواتف المحمولة) مع إحصاءات النقل في مشروع قياس كثافة المرور على الطرق السريعة – ومتابعة حوادث السير وتحديد أماكنها – وقياس عدد الأفراد العاملين أو الدارسين في موقع مختلف عن مكان الإقامة .مشروع إحصاءات النقل المنتجة من بيانات كاميرا المراقبة لتقدير حركة المرور في الطرق السريعة، الهدف الرئيسي هو دعم لتقدير حركة المرور الواردة والصادرة على حدود الدولة، واستخدام البيانات المطلوبة على المركبات لتحليل حركة المرور (جمهورية مصر العربية: الجهاز المركزي للتعبئة العامة والاحصاء).

30.4 ما الوضع الراهن لمشاريع البيانات الكبيرة؟

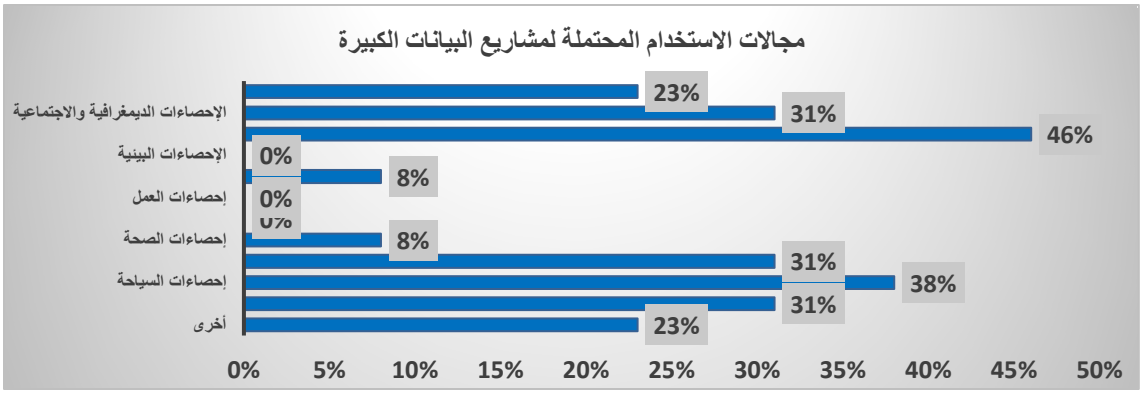
هناك حوالي 34% من مشاريع البيانات الكبيرة التي تمت الموافقة عليها ولم يتم تمويلها بعد، و33% مشاريع جارية في مرحلة التنفيذ، بينما 13% فقط اكتملت.

شكل 24. الوضع الراهن لمشاريع البيانات الكبيرة



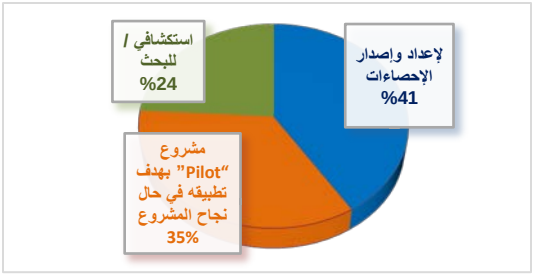
31.4 ما مجالات الاستخدام المحتملة لمشاريع البيانات الكبيرة؟

شكل 25. مجالات الاستخدام المحتملة لمشاريع البيانات الكبيرة



32.4 مشاريع البيانات الكبيرة مؤهلة كالتالي :

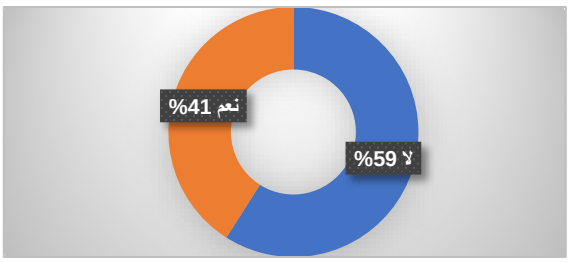
شكل 26. مشاريع البيانات الكبيرة المؤهلة



هناك حوالي 41% من مشاريع البيانات الكبيرة مؤهلة لإعداد وإصدار الاحصاءات، و35% منها مشاريع تجريبية "Pilots"، بينما 24% من المشاريع تندرج في إطار استكشافي للبحث.

33.4 أ. هل لدى مؤسستكم أي شراكات مع منظمات أو موفري بيانات آخرين حول هذا المشروع؟

شكل 27. وجود شراكات مع منظمات أو موفري بيانات آخرين حول مشروع البيانات الكبيرة



حوالي 41% من الجهات العربية المعنية تتوفر على شراكات في مجال البيانات الكبيرة مع منظمات أو موفري بيانات.

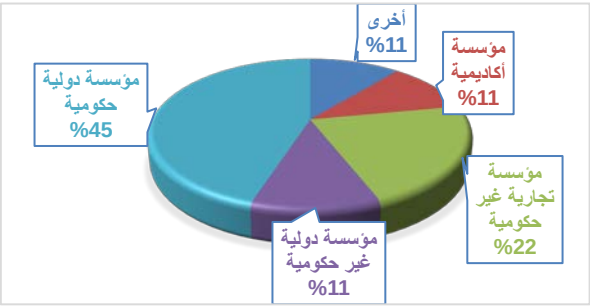
ب. اسم الشريك: تصنيف عملي

- هيئة تنظيم الاتصالات، وشركات الهاتف المحمول المحلية، والمصرف المركزي.
- هيئة تنظيم الاتصالات – الاشراف العام على الاتصالات، وشركة الاتصالات (مزود بيانات الهاتف المحمول).
- بروتوكول مع مصدر البيانات للأسعار (هايبر ماركت).
- الأمم المتحدة في إطار مبادرتها المتعلقة بالبيانات الكبيرة: " United Nations Global Pluse".
- المؤسسات الحكومية.
- هيئة الإحصاء بالأمم المتحدة، هيئة تنظيم الاتصالات، هيئة تقنية المعلومات، شركات الاتصالات، مجلس البحث العلمي، الجامعات.

ج. ما نوع المؤسسة الشريكة؟

هناك 45% من الشراكات في مجال البيانات الكبيرة مع مؤسسات دولية حكومية، و11% مع مؤسسات حكومية غير دولية، بينما 22% من الشراكات مع مؤسسات تجارية غير حكومية، و11% مع مؤسسات أكاديمية.

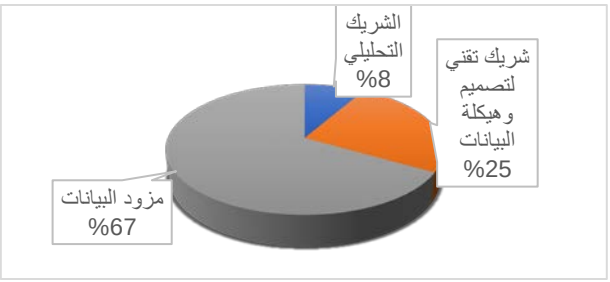
شكل 28. نوع المؤسسات الشريكة



د. ما نوع الشراكة؟

تشمل 67% من الشراكات في مجال البيانات الكبيرة شراكات مع مزودي البيانات، و25% مع شريك تقني لتصميم وهيكلة البيانات، بينما 8% مع الشريك التحليلي.

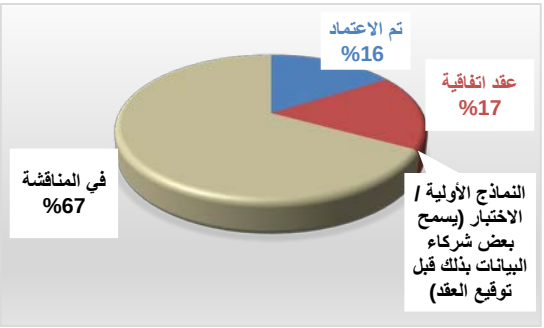
شكل 29. نوعية الشراكات



ه. ما الوضع الحالي للشراكة؟

حوالي 67% من الشراكات في مجال البيانات الكبيرة في المناقشة، و17% تم عقد اتفاقيات حولها، بينما 16% تم اعتمادها.

شكل 30. الوضع الحالي للشراكات



5. أفضل الممارسات لتطبيقات البيانات الكبيرة لأغراض الاقتصاد الكلي وفي الإحصاءات الرسمية

أ. استخدام بيانات الأعمال التجارية والبيانات الإدارية لأغراض إحصائية:

بهدف استخدام البيانات الإدارية والتجارية لأغراض إحصائية في الحسابات القومية، لذا ينبغي التحقق وفهم الخصائص الإحصائية لهذه البيانات وتحسين دقتها من خلال عملية الترجيح، وملء فجوات التغطية، وتعديلات التحيز، ومعايير المقارنة والموازنة. من بين تطبيقات بيانات الأعمال التجارية والإدارية ما يلي:

- تجريف بيانات الإنترنت (Internet scraping) لإنتاج مؤشر أسعار المستهلكين "أسعار المليار".
- استخدام بيانات الرواتب من شركة معالجة البيانات التلقائية (ADB) لعملها الشهري.
- استخدام البيانات البريدية الدولية لإنشاء مؤشر الرسالة الدولية الذي يمكن أن يُستخدم كمؤشر استرشادي رائد أو لتحسين دقة التنبؤات.
- استخدام عمليات بحث من خلال "google" لأغراض التنبؤ الآني عن حالة الاقتصاد.

ب. تطبيق البيانات الكبيرة لأغراض التنبؤ في مجالات مختلفة:

تهتم العديد من المصارف المركزية في العالم بإنشاء خريطة الطريق ومشاريع رائدة وكذا تبادل الخبرات في مجال البيانات الكبيرة، حيث يتم تطبيق هذه البيانات لأغراض التنبؤ في مجالات مختلفة، مثل: معدل البطالة، ومبيعات التجزئة، ومبيعات المنازل، والسفر والسياحة، ومبيعات وشراء السيارات، وثقة المستهلك، واتجاهات سوق الإسكان، والاستهلاك الخاص، والتضخم، وغير ذلك.

هناك العديد من الدراسات التي تطرقت لجوانب الاقتصاد الكلي باستخدام البيانات الكبيرة، نذكر منها على سبيل المثال وليس الحصر:

جدول 7: أمثلة عن الدراسات التي تطرقت لجوانب الاقتصاد الكلي باستخدام البيانات الكبيرة

مجالات استخدام البيانات الكبيرة في الاقتصاد الكلي	الدراسات
معدل البطالة، مبيعات التجزئة، مبيعات المنازل، السفر / السياحة، مبيعات السيارات، ثقة المستهلكين. Unemployment rate, retail sales, home) (sales, travel/tourism, car sales, consumer confidence	Hal Varian & Choi (2009, 2011, 2013)
معدل البطالة في ألمانيا DE unemployment rate	Zimmermann K & Askitas N (2009)
معدل البطالة بالولايات المتحدة US unemployment rate	D'Amuri F, & Marcucci J (2010, 2013)
معدل البطالة وسوق الإسكان في المملكة المتحدة UK unemployment rate & housing market trends	McLaren N & Shanbhogue R (2011)
الإستهلاك الخاص في ألمانيا DE private consumption	Vosen & Schmidt (2011)

تطبيقات البيانات الكبيرة في الدول العربية	
Carriere-Swallow (2011)	Car purchases in Chile شراء السيارات في دولة الشيلي
Guzmán G (2011)	Inflation التضخم
Fantazzini D & Toktamysova Z (2014)	German car sales شراء السيارات في ألمانيا
Morgan J, e all (2015)	DE, FR, IT, ES NL معدلات البطالة
Simeon Vosen and Torsten Schmidt (2011)	Forecasting private consumption: survey-based indicators vs. Google trends
A. Cavallo and R. Rigobon (2016)	The Billion Prices Project: Using Online Prices for Measurement and Research
J. de Haan, L. Willenborg and A.G. Chessa (2016)	An overview of price index methods for scanner data

كما يمكن إيجاد المزيد من مراجع الدراسات المتعلقة بموضوع البيانات الكبيرة في تقرير صندوق النقد الدولي حول "Big Data: Potentials, Challenges and Statistical Implications 2017".

ج. تجربة فريق العمل العالمي للأمم المتحدة المعني بالبيانات الكبيرة (UN-GWG):

وافقت اللجنة الإحصائية للأمم المتحدة في دورتها الخامسة والأربعين (في نيويورك 2014) على إنشاء فريق العمل العالمي المعني بالبيانات الكبيرة لأغراض الإحصاءات الرسمية الذي يهدف إلى متابعة فوائد وتحديات البيانات الكبيرة، وإمكانات البيانات الكبيرة لرصد أهداف التنمية المستدامة والإبلاغ عنها.

يتكون الفريق من مجموعات (task teams) مكلفة بمجالات محددة:

- الوصول إلى مصادر البيانات الكبيرة والشراكات.
- البيانات الكبيرة وأهداف التنمية المستدامة.
- بيانات الهاتف الجوال.
- صور الأقمار الصناعية والبيانات الجغرافية المكانية.
- بيانات وسائل الإعلام الإجتماعية.
- التدريب والمهارات وبناء القدرات.
- لجنة المنتدى العالمي للبيانات والخدمات والتطبيقات.

جدول 8: الإنجازات والتقدم المحرز لفريق العمل العالمي للأمم المتحدة المعني بالبيانات الكبيرة

2014	2015	2016	2017
- إنشاء فريق العمل العالمي الأمم المتحدة المعني بالبيانات الكبيرة (UN-GWG). - تقاسم "Pilots" حول مشاريع البيانات الكبيرة. المؤتمر الأول: في الصين	- الانتقال من "Pilots" إلى المبادئ التوجيهية. المؤتمر الثاني: في الإمارات العربية المتحدة	- الوصول إلى مصادر والشراكات البيانات الكبيرة، وأهداف التنمية المستدامة المؤتمر الثالث: في إيرلندا	- التعاون في مجال البيانات والبيانات الموثوقة (Data collaboratives and trusted data) المؤتمر الرابع: في كولومبيا

تطوير مخزونات البيانات الكبيرة
(Big Data Inventory)
https://unstats.un.org/bigdata/inventor

تطوير بوابة الويب على البيانات الكبيرة
(Web gate on big data)
https://unstats.un.org/bigdata

د. تجربة البنك المركزي الأوروبي:

تشمل تجربة البنك المركزي الأوروبي في مجال البيانات الكبيرة العديد من المشاريع الرائدة وأوراق العمل وكذلك التطبيقات العملية لأغراض الإقتصاد الكلي باستخدام البيانات الكبيرة.

- **من أهم المشاريع "Pilots" ذات الصلة:**
  - مجموعة البيانات الإدارية (على سبيل المثال بيانات الميزانية العمومية للشركات).
  - مجموعة بيانات بحث عبر الشبكة الإلكترونية (مثل معلومات البحث عبر "Google").
  - مجموعة البيانات التجارية (مثل مشغلي بطاقات الائتمان).
  - بيانات السوق المالية (على سبيل المثال بيانات التداول عالية التردد).

- **أوراق إحصائية مرتبطة باستخدام البيانات الكبيرة:**

- "التنبؤ الآني للناتج المحلي الإجمالي باستخدام بيانات المدفوعات الإلكترونية"، حيث يمكن استخدام معاملات الدفع الإلكترونية في عملية التنبؤ الآني بنمو الناتج المحلي الإجمالي. إضافة إلى أن معاملات بطاقات الخصم تساهم في تحسين دقة التنبؤ.
- "العلاقات بين التغيرات في ثقة المستهلك ووسائل الإعلام الإجتماعية العامة في هولندا"، وإمكانية استخدام وسائل الإعلام الإجتماعية كمؤشر للتغيرات في ثقة المستهلك وكمؤشر مبكر.
- تحديد آثار ارتفاع معاملات الأسواق المالية الدولية "online bullishness" عبر الشبكة الإلكترونية، حيث تم تطوير تقدير لتوجهات المستثمرين (investors sentiment)، استناداً إلى طلبات البحث على "Twitter" و"Google"، إضافة إلى بيان الإرتباط الإيجابي بين كل منهما وبين ثقة المستثمرين.
- أحد أهم الدراسات ذات الصلة التي قام بها البنك المركزي الأوروبي تشمل كيفية استخدام بيانات البحث "Google" للتنبؤ الآني بمعدلات البطالة في منطقة اليورو، حيث تم استخدام هذه البيانات من تغييرات الحجم الأسبوعية لمؤشرات البحث عبر "Google".

ه. تجربة مكتب الاحصاء الاوروبي (EUROSTAT):

تعتبر تجربة مكتب الاحصاء الاوروبي في مجال البيانات الكبيرة غنية، لما تشمله من مشاريع مختلفة لإستكشاف مختلف مصادر البيانات الكبيرة وأيضاً بنية تقنية المعلومات والشراكات، إلى جانب وضع مبادئ توجيهية وأطر عامة، كذلك إنشاء ورش عمل مع مقدمي البيانات والبحوث والمنظمات الدولية، إضافة إلى التعاون مع الأمم المتحدة فيما يخص إطار العمل المنهجي.

من أولويات المكتب دمج الإحصاءات الرسمية في استراتيجيات البيانات الكبيرة، وضمان الحصول على البيانات واستمرارية الوصول إليها، إضافة إلى معالجة التحديات المرتبطة بجودة وأمن وخصوصية البيانات (data security & privacy concerns).

نذكر من أهم المشاريع "Pilots" التي تم توقيع اتفاقات شراكة حولها، ما يلي:

- التجريف على الشبكة الإلكترونية (Web scraping)، وتشمل الوظائف الشاغرة وخصائص

المؤسسات.
- العدادات الذكية في مجال استهلاك الكهرباء والمساكن المؤقتة.
- بيانات للتعرف على أنظمة السفن.
- بيانات الهاتف الجوال: على سبيل المثال مشروع دراسة الجدوى بشأن استخدام بيانات تحديد المواقع المتنقلة في الإحصاءات السياحية *the use of mobile) positioning data for tourism statistics*).

نظراً للتحديات المرتبطة بجودة البيانات والأخلاق والاتصالات، يعمل المكتب على مشروع تحليل التشريعات والاستراتيجيات، كذلك الأخلاقيات والاتصال *(ethics and communication)*.
من جهة أخرى، يوفر مكتب الاحصاء الأوروبي برنامج تدريبي للإحصائيين الأوروبيين، يشمل:
- دورات مخصصة للبيانات الكبيرة.
- التركيز على مصادر وأدوات البيانات الكبيرة.
- اكتساب المهارات اللازمة لتقييم المصادر ونوعيتها، والمهارات اللازمة لاستخدام الأدوات، واستكشاف مصادر البيانات الكبيرة.

قام مكتب الاحصاء الاوروبي "أوروستات" بمطابقة سجل الأعمال الحالي (للقطاع العام والخاص) مع البيانات التجارية على المستوى الجزئي والإجمالي لإنشاء بيانات تفصيلية عن خصائص شركات التصدير والاستيراد، مما يمكن أن يساعد في تصميم سياسات التجارة والاستثمار بشكل كبير.

كما يمكن استغلال البيانات التي يتم تجميعها لأغراض تجارية وإدارية لأغراض أخرى، فبالنسبة للأعمال التجارية، تشمل الأغراض الأخرى التسويق وخطط الموقع الجغرافي وخطط الاستثمار القصيرة والطويلة الأجل وقياس أداء الشركة على مستوى الصناعة. أما بالنسبة للقطاع العام، تشمل هذه الأغراض الأخرى التخطيط الجغرافي للبنية التحتية وتوفير الخدمات، لفهم الاستجابات السلوكية وخصائص السكان لتصميم التأمين الصحي والبطالة، وكذلك للسياسات الضريبية وغيرها من السياسات، لتتبع حدوث المرض لأغراض الصحة العامة، ولتقييم فعالية البرامج الحكومية.

و. تجربة الحكومة الصينية لتطوير "نظام الائتمان الاجتماعي" باستخدام البيانات الكبيرة:

وضعت الحكومة الصينية خططاً لتصنيف كل واحد من مواطنيها حسب درجة ائتمان باستخدام البيانات الكبيرة في إطار مبادرة تطوير "نظام الائتمان الاجتماعي" لكل مواطن، استناداً إلى البيانات الحكومية المتعلقة بوضعه الإقتصادي والإجتماعي واستخدام البيانات الكبيرة.
ترتكز هذه المبادرة على استغلال قوة البيانات الكبيرة بما فيها بيانات الهواتف الذكية والتجارة الإلكترونية ووسائل الإعلام الإجتماعية لربط "نظام الإئتمان الإجتماعي" الجديد بحوالي 1.3 مليار بطاقة هوية وطنية من المواطنين الصينيين، وتعيين تصنيف للمواطنين حسب سلوكهم و"أنشطة الأصدقاء لديهم في الرسم البياني الاجتماعي الخاص بهم من خلال وسائل الإعلام الإجتماعية".
من خصائص هذا النظام، لأنه يمكن لدرجة الإئتمان لكل مواطن أن تتأثر بشكل سلبي بهوايته وعاداته التسوقية ونمط حياته، وكذلك ما ينشره على الشبكة الإلكترونية من تعليقات وآراء سياسية، إضافة إلى ما يقوم به من تواصل اجتماعي، مثل الشراء أو النشر أو القراءة.

كما تشمل الخطة أيضاً استخدام بيانات السلطات القضائية (المحاكم) والشرطة والبنوك والسجلات الضريبية وبيانات التوظيف. إضافة إلى ذلك، يمكن أن يتم تصنيف الأطباء والمعلمين والبلديات والشركات من قبل المواطنين حسب نزاهتهم ومستوياتهم الاحترافية.
بحلول عام 2020، سوف تكون درجة الإئتمان الصينية الجديدة إلزامية، لكن يمكن للمواطنين حالياً تتبع درجات تصنيفهم عن طريق التطبيق الحر"Sesame Credit".

ز. تجربة مكتب الإحصاءات الأسترالي:

يمكن للأجهزة الاحصائية في الدول العربية الاستفادة من خبرة مكتب "ABS's international" الأسترالي، المكلف من قبل مكتب الإحصاءات الأسترالي بالتعاون الاستراتيجي مع الجهات المختصة على الصعيد الدولي، لتعزيز التعاون في مجالات الإحصاءات الناشئة، واستخدام مصادر البيانات الإدارية لأغراض إحصائية، إضافة إلى تعميم الاستفادة من إطار قياس الإقتصاد والمجتمع القائمان على المعرفة الذي تعمل عليه الحكومة الأسترالية، لتطوير مؤشرات العلم والتقنية، وإدماج المعلومات الإحصائية والجغرافية المكانية.

ح. تجارب أخرى:

- **صندوق النقد الدولي**: ركز صندوق النقد الدولي في تقريره 2017 حول " Big Data: Potentials, Challenges and Statistical Implications" على أهمية البيانات الكبيرة في معالجة التأخير في الفترات الزمنية للإحصاءات الرسمية، وكذلك في دعم عملية التنبؤات للمؤشرات القائمة، حيث تم التركيز على دور البيانات الكبيرة كمصدر للبيانات والابتكار في إنتاج الإحصاءات الرسمية.
- **منظمة الأغذية والزراعة (FAO)**: تقوم المنظمة بعدد من الدراسات حول استخدام البيانات الكبيرة في مجال أسعار المواد الغذائية، نذكر منها دراسة حول "التنبؤ الآني لأسعار استهلاك المواد الغذائية باستخدام البيانات الكبيرة: تكامل القطاعين العام والخاص".

ط. أمثلة عن بعض البحوث وعمليات التحليلات التعاونية الأولية:

- دراسة التغيرات المتعلقة بأسعار المواد الغذائية وتضخم أسعار المواد الغذائية، وإمكانية استخدام وسائل الإعلام الإجتماعية بيانات "Facebook" و"Twitter" لأغراض الإحصاءات الرسمية، مع التحقق من صحة، وطرق تنقيح ومعالجة البيانات التي تم تجميعها من خلال وسائل الاعلام الإجتماعية، ووضع تقديرات متينة (robust) لاستخدامها مع الإحصاءات الرسمية.
- دراسة تقييم جدوى استخدام بيانات تحديد المواقع المتنقلة لقياس السياحة، التي قام بها مكتب الإحصاء الأوروبي (Eurostat)، حيث أشارت الدراسة إلى إمكانيات البيانات الكبيرة، وأكدت في نفس الوقت على ضرورة حل عدد من القضايا من بينها: خصوصية البيانات والقضايا التنظيمية المرتبطة بالخصوصية، والرأي العام المتعلق باستخدام بيانات الهواتف المتنقلة، والحواجز المالية والتجارية التي تعوق الوصول إلى بيانات الهواتف النقالة، إضافة إلى التحديات التقنية لاستخدام هذه البيانات والوصول إليها، كذلك تحديات تطبيق المنهجيات التي تمكن من الاستخدام الفعال لهذه البيانات.

- دراسة العلاقة بين اتجاهات المستهلكين الشهرية "sentiments" من خلال وسائل الإعلام الإجتماعية ("Facebook" ورسائل "Twitter") وحالة الإقتصاد في هولندا. بينت الدراسة وجود ارتباط عالي على مدى السنوات الثلاث والنصف التي تمت فيها الدراسة (2010-2013)، وهو مثال جيد للكيفية التي يمكن بها للبيانات الكبيرة أن تكمل، أو أن تكون متكاملة مع المؤشرات الحالية لحالة الاقتصاد.
- دراسة "Anson and Heble" لكيفية استخدام البيانات البريدية لوضع مؤشر الرسائل البريدية. وعلى الرغم من تراجع استخدام الخدمة البريدية في جميع أنحاء العالم، بينت الدراسة أن المؤشر يعمل بشكل جيد كمؤشر رئيس يكشف تغيرات الاتجاه قبل المؤشر التجاري على مدى السنتين لفترة الدراسة (2010-2012)، فضلاً عن بيان الاختلافات في الأداء بين البلدان في أوروبا. كما أشارت الدراسة أيضاً إلى أن استخدام البيانات البريدية التي تعكس الوقت الحقيقي، يمكن أن يحسن دقة التنبؤ، مع مراعاة أن استخدامها في إطار مؤشرات أو تنبؤات رئيسية، يتطلب تطبيق اختبارات على مدى فترة زمنية أطول، وتطوير أساليب لا تستكشف فقط اتجاه التغييرات، بل الحجم النسبي للتغيرات، وتطبيق اختبارات عبر البلدان.

ي. الاستفادة من التعاون بين الوكالات الإحصائية الرسمية وشركات المعلومات:

جاءت بعض أنجح الممارسات المتعلقة باستخدام "البيانات الكبيرة" من خلال بناء شراكات والعمل على مشاريع مشتركة تجمع مزودي البيانات الخاصة مثل "Google"، و"Baidu"، و"Yahoo"، والوكالات الإحصائية الرسمية. توفر هذه المشاريع المشتركة فوائد متبادلة هامة، من بينها تجربة المكتب الأمريكي للتحليل الاقتصادي في التعاون مع شركات المعلومات.

ك. تجربة المكتب الأمريكي للتحليل الاقتصادي في التعاون مع شركات المعلومات:

شملت تجربة المكتب الأمريكي للتحليل الاقتصادي العديد من مجالات التعاون مع شركات المعلومات، على رأسها التعاون مع شركة "IBM" لتطوير أسعار الحواسيب الهائلة، كذلك التعاون مع شركة "Chrysler" ومصنعي السيارات الآخرين لتطوير بيانات تسعير السيارات والشاحنات الجديدة، إضافة إلى بناء شراكات مع "Google" لتطوير محركات البحث الإقليمية.

ل. بعض الأدوات المستخدمة في مجال البيانات الكبيرة:

جدول 9: بعض الأدوات المستخدمة في مجال البيانات الكبيرة (المصدر: الشبكة الإلكترونية)

Data Cleaning معالجة البيانات OpenRefine – DataCleaner	**تخزين البيانات وإدارتها** Hadoop - Cloudera - MongoDB - Talend
Data Integration دمج البيانات Blockspring - Pentaho	**Data Collection تجميع البيانات** Import.io
Data Mining تنقيب البيانات RapidMiner - IBM SPSS Modeler - Oracle data mining - Teradata – FramedData – Kaggle	**Data Languages** R Language – Python – RegEx – Xpath
Data Visualization العرض المرئي للمعلومات Tableau – Silk – CartoDB – Chartio - Plot.ly – Datawrapper	**Data Analysis تحليل البيانات** Qubole – BigML – Statwing

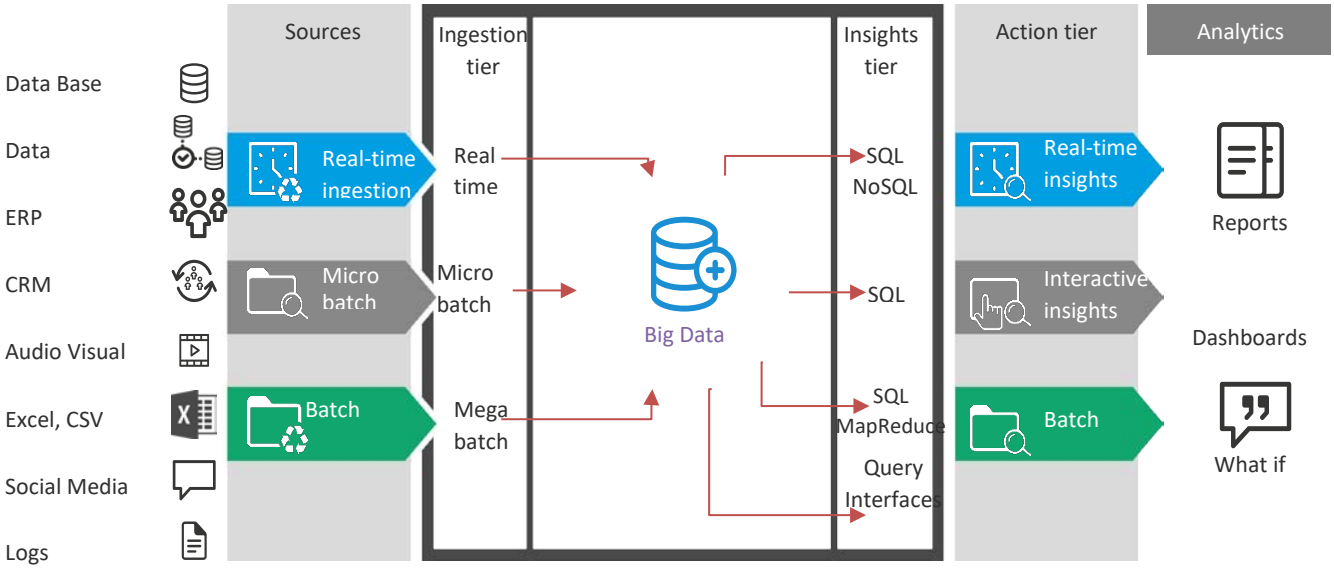
م. نموذج مشروع البيانات الكبيرة: مركز البيانات أو بحيرة البيانات

تعتبر بحيرة البيانات "Data Lake or Data Hub" مستودع تخزين يحمل كمية هائلة من البيانات الخام في شكلها الأصلي (حتى تكون الحاجة إلى استخدامها ومن ثم هيكلتها). على عكس بحيرة البيانات، يمكن لقاعدة البيانات ومستودع البيانات تخزين البيانات التي خضعت فقط لهيكلة وتنظيم مُسبق. مع تزايد حجم ونوع البيانات، أصبحت بحيرة البيانات مرادفاً للبيانات الكبيرة، حيث توفر طرقاً لمواجهة تحديات البيانات الكبيرة بتكلفة منخفضة وممكنة من الناحية التقنية، كما يتزايد الطلب على استخدام بحيرة البيانات والانتقال من اعتماد بنية تصميم تقليدية.

تختلف بحيرة البيانات ومستودع البيانات (Datawarehouse) من حيث العديد من الجوانب وهي: البيانات، والتجهيز، والتخزين، وخفة الحركة، والأمن والمستخدمين. فيما يلي الاختلافات بين المصطلحين:

- **البيانات:** على عكس قاعدة البيانات ومستودع البيانات التي تقوم بتخزين البيانات التي تم تنظيمها فقط، فإنه يمكن لبحيرة البيانات تخزين جميع أنواع البيانات سواء كان مهيكلة (structured)، شبه مهيكلة، أو غيرمهيكلة ".
- **المعالجة:** قبل أن يتم تحميل البيانات في مستودع بيانات، ينبغي أن يعطى لها أولاً شكلاً وهيكلاً. وبعبارة بينما تقبل بحيرة البيانات، البيانات في شكلها الخام، ويتم تزويدها بشكل وهيكل عندما تكون الحاجة إلى استخدامها.
- **التخزين:** من أحد السمات الرئيسية لتقنية البيانات الكبيرة هو النظر في تكلفة تخزين البيانات. يعتبر تخزين البيانات باستخدام تقنية البيانات الكبيرة أرخص نسبياً مقارنة بتخزين البيانات في مستودع البيانات، بينما يمكن أن يكون تخزين مستودع البيانات مكلفاً، خاصة إذا كان حجم البيانات كبير. كما أن بحيرة البيانات مصممة لتخزين المعلومات بتكلفة منخفضة.
- **أمن المعلومات:** تعد تقنيات البيانات الكبيرة التي تعتمد على استخدام بحيرات البيانات جديدة نسبياً، لذا، فإن قدرتها على ضمان أمن البيانات أقل نضجاً مما عليه الحال بالنسبة لمستودع البيانات.
- **المستخدمون:** تختلف مستودعات البيانات، وبحيرات البيانات وقواعد البيانات من حيث المستخدمين، تستخدم مستودعات البيانات في الغالب في قطاع الأعمال من قبل رجال الأعمال والمهنيين، بينما تستخدم بحيرات البيانات في الغالب في المجالات العلمية من قبل علماء البيانات، في حين أن قواعد البيانات مرنة جداً وهي مناسبة لأي مستخدم. يوضح الشكل رقم (31) مركز البيانات أو بحيرة البيانات.

شكل 31. بحيرة أو مركز البيانات "Data Lake or Data Hub"



المصدر: الاتحاد الدولي للإتصالات (International Telecommunication Union)

ن. التقارب بين البيانات الكبيرة، وإنترنت الأشياء، وتحليل ومعالجة البيانات "at the Edge":

مع تزايد تدفقات البيانات و"إنترنت الأشياء" يطرح التساؤل نفسه "*هل ستكون تحليلات البيانات التي تتم بالقرب من مصدرها، ما يطلق عليه مسمى "at the Edge" هي الوضع المستقبلي للبيانات الكبيرة"؟*

تربط إنترنت الأشياء موجة جديدة من الأجهزة، مثل أجهزة الاستشعار الصناعية، والأجهزة التي يمكن ارتداؤها، وكاميرات البيع بالتجزئة، وشبكة الإنترنت وغيرها. يمكن للبيانات التي يتم تجميعها من هذه الأجهزة أن تساعد على خفض التكاليف، وزيادة الإيرادات، وتحسين خدمة العملاء، فقط في حال إمكانية استغلال وتحليل هذه البيانات على وجه السرعة.

يوفر نهج "إنترنت الأشياء" العديد من الفوائد، على رأسها إمكانية الوصول إلى المزيد من البيانات وتوفير رؤى أفضل. كما أنه يخلق تحديات جديدة، حيث أن الحجم الهائل وسرعة البيانات التي يتم جمعها من الأجهزة عبر الإنترنت، يمكن أن يجهد البنية التحتية للشبكة، وأن يسبب في تأخير معالجة وتحليل البيانات نتيجة ازدحام النطاق الترددي لتدفقات البيانات أو توجيهها غير الفعال.
لحل هذه القضايا، يكتسب مفهوم "Edge computing" أهمية كبيرة، لكونه نهج تحليلي، يساعد على تحليل البيانات بالقرب من مصدرها ("at the edge")، يعني أساساً تصميم النظم حيث يتم تنفيذ تحليل البيانات عند النقطة التي (أو قريبة جداً من حيث) يتم جمع هذه البيانات منها، بدلاً من إرسالها إلى خادم بعيد " remote server" لتحليلها باستخدام الحوسبة السحابية "cloud computing".

تساعد التحليلات التي تتم بالقرب من مصدرها "at the Edge"، المؤسسات (أو حتى الأجهزة نفسها) على العمل على رؤى جديدة في غضون ميلي ثانية بدلاً من الانتظار لنقل البيانات إلى مستودع البيانات المركزية

للمعالجة بعد بضع ساعات. حسب تقرير"IDC FutureScape for Internet of Things"، تعتبر معظم الشركات أنه بحلول عام 2018، حوالي 40 بالمائة من البيانات التي تولدها بواسطة "إنترنت الأشياء" سيتم تخزينها ومعالجتها بالقرب من مصدرها، وليس في مستودعات بيانات مركزية.

مع النمو السريع لمجموعات البيانات والأجهزة التي تدعم "إنترنت الأشياء" أكثر من أي وقت مضى، فمن المرجح أن يصبح التوجه الاستراتيجي نحو مبدأ التحليل بالقرب من المصدر"Analytics on the Edge"، متزايد الأهمية لتنفيذ بنية تحليلات أكثر كفاءة، الأمر الذي سيقرب التحليل من البيانات، وليس العكس.

6. مقترحات وضع إطار لبرنامج عمل إقليمي لدعم تطبيق البيانات الكبيرة في الدول العربية

تتضمن مقترحات وضع إطار لبرنامج عمل إقليمي لدعم تطبيق البيانات الكبيرة في الدول العربية ثلاث محاور رئيسية:

المحور الأول: إنشاء فريق عمل إقليمي معني بالبيانات الكبيرة أو الإستفادة من الأطر القائمة حالياً.

المحور الثاني: مبادرة إقليمية لبناء القدرات في مجال البيانات الكبيرة.

المحور الثالث: إنشاء شبكة إقليمية للتعاون وبناء الشراكات في مجال البيانات الكبيرة.

1.6 المحور الأول - إنشاء فريق عمل إقليمي معني بالبيانات الكبيرة

يمكن أن يتكون فريق العمل الإقليمي المعني بالبيانات الكبيرة تحت مظلة مبادرة "عربستات". وتشمل عضويته خبراء من المصارف المركزية ومؤسسات النقد العربية، وأجهزة الإحصاء في الدول العربية، ووزارات المالية ويتولى صندوق النقد العربي دور السكرتارية الفنية للفريق. كما يمكن أن يضم مؤسسات دولية من جهات مثل منظمة التعاون والتنمية الاقتصادية، ولجنة الأمم المتحدة الاقتصادية (United Nations Economic Commission)، إضافة إلى صندوق النقد والبنك الدوليين، وخبراء أكاديميين.

من مهام فريق العمل المعني، العمل على تنفيذ مخططات البرنامج الإقليمي لتلبية احتياجات الدول الاعضاء في مجال البيانات الكبيرة. يشمل الإطار العام لبرنامج العمل الإقليمي عدداً من المجالات المترابطة بعضها البعض، وهي:

أ. التعاون وتبادل أفضل الممارسات والمنهجية والانتقال إلى "الاستخدام الحقيقي" للبيانات الكبيرة والتعرف على مختلف مصادرها.

ب. العوامل الرئيسة التي تمكن من تطبيق البيانات الكبيرة بشكل ناجح، والمهارات والبنية التحتية لتقنية المعلومات.

ج. السياسات والإطار التنظيمي.

د. إطار جودة البيانات الكبيرة و"الأخلاق والاتصال".

ه. الحوكمة.

شكل 32: الإطار العام لبرنامج عمل إقليمي لدعم تطبيق البيانات الكبيرة في الدول العربية

أ. التعاون وتبادل أفضل الممارسات والمنهجيات والانتقال إلى "الاستخدام الحقيقي" للبيانات الكبيرة.

يمكن معالجة قضايا التعاون وتبادل الخبرات وتعزيز المعرفة للمنهجيات المستخدمة، وكذا الإنتقال إلى "الاستخدام الحقيقي" للبيانات الكبيرة، من خلال ما يلي:

- **إطلاق سلسلة من المشاريع "Pilots"**. تعد المشاريع "Pilots"، من أهم محاور خارطة الطريق وخطة العمل التي يمكن أن يقوم بها فريق العمل على مدى سنوات. كما يمكن لهذه المشاريع أن تكون ركيزة هامة من أنشطة البيانات الكبيرة في المعيار البيئي والإجتماعي في السنوات المقبلة، وينبغي أن تمهد الطريق نحو توليد منتجات إحصائية وغيرها استناداً لاستخدام البيانات الكبيرة.
- **إقامة تعاون وثيق** بين فريق العمل المعني بالبيانات الكبيرة في الدول العربية وفريق عمل الأمم المتحدة المعني بالبيانات الكبيرة " UN Global Working Group on Big Data for Official Statistics"، من أجل تجنب العمل المزدوج وضمان التعاون بين المجموعتين.
- **تبادل الخبرات:** اتخاذ مبادرات تعاون وثيق من خلال صندوق النقد العربي بين فريق العمل الإقليمي المعني بالبيانات الكبيرة المقترح تشكيله، مع منظمات دولية مختصة لأغراض الإحصاءات الرسمية والتنبؤ الإقتصادي، وكذلك التعاون مع لجنة "Irving Fisher on Central Bank Statistics" حول إحصاءات البنوك المركزية، الذي يعمل تحت رعاية بنك التسويات الدولية، ويشمل اقتصاديين وإحصائيين في البنوك المركزية، فضلاً عن خبراء آخرين لمناقشة القضايا الإحصائية التي تهم البنوك المركزية.

ب. العوامل الرئيسة التي تُمكن من تطبيق البيانات الكبيرة بشكل ناجح: المهارات والبنية التحتية لتكنولوجيا المعلومات.

هناك حاجة إلى مهارات جديدة وهندسة تقنية المعلومات والبنية التحتية، والحاجة كذلك إلى التكيف مع مصادر بيانات جديدة. يتمثل أحد الإجراءات الملموسة في إعداد سلسلة من الدورات التدريبية في معهد

السياسات الإقتصادية بصندوق النقد العربي لتطوير تطبيقات البيانات الكبيرة. وستكون أهم المهام المقترحة لفريق العمل الاقليمي المعني بالبيانات الكبيرة المقترح تشكيله، إعداد الخطوط العريضة لتحسين المهارات والبنية التحتية لتقنية المعلومات، وتركيز الدورات التدريبية على مصادر البيانات والأدوات التي يمكن استخدامها.

- هناك مجموعة من المهارات والمعرفة يتطلب توفرها لأغراض معالجة البيانات الكبيرة، والتي لا يمكن تغطيتها بشكل جيد في دورة واحدة. لذا، هناك حاجة إلى عدة دورات لتغطية هذه الاحتياجات.
- هناك مهارات مهمة لا تتعلق فقط بالبيانات الكبيرة يتطلب توفرها بهدف العمل مع مصادر البيانات الكبيرة مثل التعلم الآلي (machine learning)، يمكن أن تغطيها الدورات التدريبية المنهجية، التي يقدمها برنامج المهارات والبنية التحتية لتقنية المعلومات.
- كما توفر الدورات التدريبية المنهجية أيضاً المهارات الهامة الضرورية لإنتاج المخرجات الإحصائية التي تعتمد على إمكانات مصادر البيانات الكبيرة، ولا سيما أساليب التنبؤ الآني (nowcasting).

ج. **السياسات والإطار التنظيمي:** يتمثل أحد أهم جوانب السياسات في دمج الإحصاءات (الرسمية) في أي استراتيجية تتعلق بالبيانات الكبيرة، نظراً لأهمية ذلك في وضع الإحصاءات على الخريطة وفتح الأبواب أمام الوصول الفعلي للبيانات. فالبيانات الكبيرة غالباً ما تحتفظ بها أو تخزنها شركات خاصة، ولا ينبغي أن يقتصر الأمر على الوصول إلى البيانات، بل ينبغي أن يشمل رؤية طويلة الأجل، أي استمرارية النفاذ (continuity of access)، وهو شرط أساسي لنظام إحصائي سليم يقوم، كلياً أو جزئياً، على البيانات الكبيرة كمصدر للبيانات.

يعد أمن وخصوصية البيانات من التحديات الرئيسة التي تحول دون الوصول إلى البيانات. كما أن هناك تحدي هام آخر يتمثل في إيجاد نموذج (business model) مستدام للبيانات الكبيرة في الإحصاءات الرسمية، مع مراعاة أثر الميزانية بالنسبة لأجهزة الإحصاء وغيرها من الجهات المعنية. لمعالجة هذه القضايا، يمكن لصندوق النقد العربي القيام بتحليل الأطر القانونية على مستوى الدول العربية.

ه. **إطار جودة البيانات و"الأخلاقيات والاتصال":**

ستظهر قضايا جودة البيانات بشكل أكبر خلال المشاريع "pilots"، عند تقييم الوصول إلى البيانات، ومن أهم القضايا المفاهيمية، على سبيل المثال كيفية الحفاظ على التعاريف الإحصائية عند استخدام البيانات الكبيرة، وكذلك التوقيت المناسب والمرونة في الوصول إلى البيانات، إضافة إلى قضايا التغطية وأخذ العينات (sampling)، وغيرها. يمكن لصندوق النقد العربي المساهمة في المبادرات المعنية بالبيانات الكبيرة بشأن تعزيز إطار جودة البيانات الكبيرة.

من ناحية أخرى، سوف تلعب "الأخلاقيات والاتصال" دوراً هاماً بل حاسماً، حيث سيحتاج صانعي السياسات والشركات إلى دعم "الرأي العام" من أجل إطلاق مبادرات او التعاون في مجال البيانات الكبيرة. كما ستصبح حماية البيانات أكثر أهمية مما هي عليه الآن.

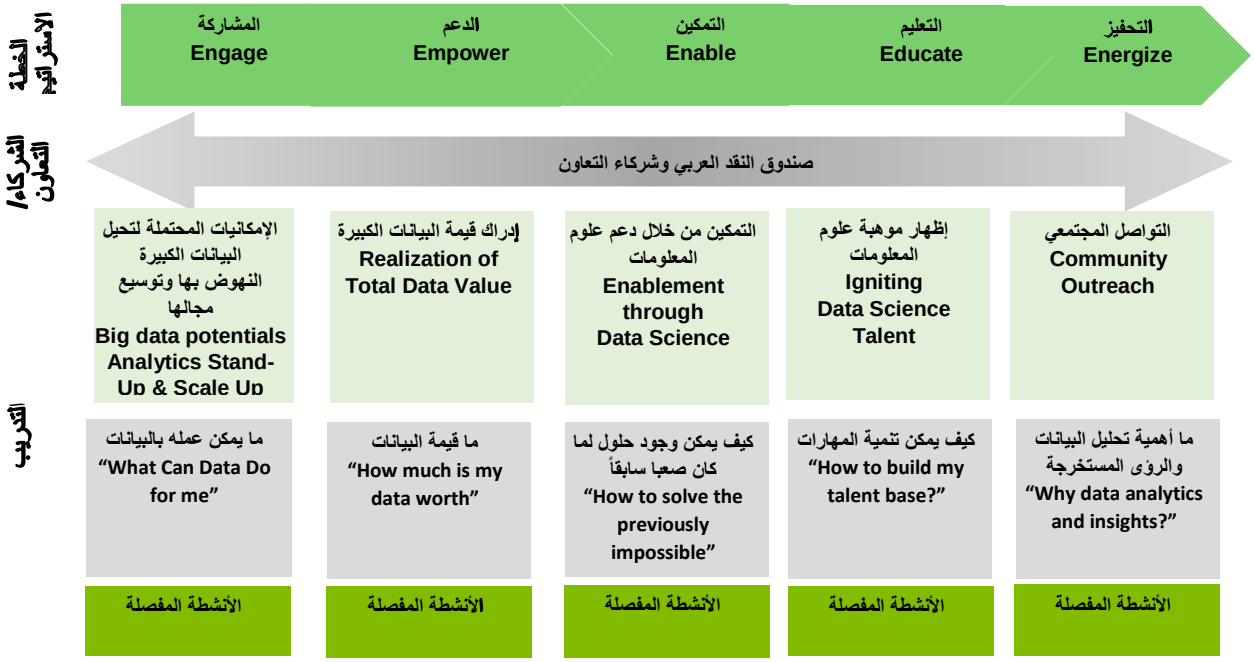
د. **الحوكمة:** المساهمة في دعم توليد إحصاءات اقتصادية عالية الجودة من خلال إطار حوكمة شامل.

2.6 المحور الثاني - مبادرة إقليمية لبناء القدرات في مجال البيانات الكبيرة

يمكن لصندوق النقد العربي إعلان نشاطه فيما يخص قيادة مبادرات استخدام البيانات الكبيرة وعمليات التحليل لأغراض الإحصاءات الرسمية وعمليات التنبؤ الاقتصادي بالمنطقة العربية، وما يعقب ذلك من برنامج تدريبي متكامل للتحليل الاحصائي واستخدام البيانات الكبيرة وخطة عمل لبناء القدرات بالتعاون مع جهات إقليمية ودولية.

كما يمكن تعزيز هذا العمل تحت مظلة مبادرة "عربستات"، والعمل على تلبية احتياجات النظم الإحصائية ووزارات المالية والمصارف المركزية في الدول العربية، من مهارات وتقنيات تستلزمها البيانات الكبيرة، لتلبية احتياجات الجهات المعنية في الدول العربية ومساعدتها في مجال استخدام هذه البيانات، كما يوضح الشكل رقم 31.

شكل 33. مبادرة إقليمية لبناء القدرات في مجال البيانات الكبيرة تحت مظلة مبادرة "عربستات"



يمكن أن تشمل المبادرة الإقليمية لدعم استخدام البيانات الكبيرة والتحليل الاحصائي والتنبؤات الاقتصادية، المحاور التالية:

أ. التعاون مع الأطر الدولية المعنية (على سبيل المثال فريق عمل الأمم المتحدة المعني بالبيانات الكبيرة (UN- GWG))، في إطار مبادرات إقليمية ودولية لتطوير أساليب وأدوات وتطبيقات البيانات الكبيرة، بما في ذلك عملية التنسيق وتسهيل عمل الفريق في المنطقة العربية.

ب. تقديم المشورة بشأن وضع برنامج للتدريب يعالج الثغرات التي يتم تحديدها في تحليل احتياجات الدول الأعضاء استناداً إلى نتائج استبيان صندوق النقد العربي حول تطبيقات البيانات الكبيرة في الدول العربية.

ج. العمل على تعزيز التعاون وتبادل الخبرات وبناء الشراكات من خلال إنشاء شبكة إقليمية للتعاون وتبادل الخبرات والمعرفة.

د. تنسيق خطط التدريب بما يتماشى مع نتائج استبيان صندوق النقد العربي وكذلك الدراسة الاستقصائية العالمية بشأن البيانات الكبيرة لأغراض الإحصاءات الرسمية، التي تمت صياغتها ضمن فريق عمل الأمم المتحدة المعني بالبيانات الكبيرة. يمكن أن تشمل الدورة التدريبية المجالات التالية حسب المستويات المختلفة التي يمكن تضمينها:

المجالات	مستوى المبتدئين	المستوى المتوسط	المستوى المتقدم
مجال البيانات			
المجال التقني			
المجال غير التقني			

ه. متابعة التطورات الدولية في مجال استخدام البيانات الكبيرة والتدريب وبناء القدرات والاستفادة من الفرص التي تتاح لخدمة برنامج العمل الإقليمي والدول الأعضاء في مجال البيانات الكبيرة.

3.6 المحور الثالث - إنشاء شبكة إقليمية للتعاون وبناء الشراكات في مجال البيانات الكبيرة

في إطار مشاركة المعرفة وتبادل الخبرات، يمكن إنشاء شبكة إقليمية للتعاون وبناء الشراكات في مجال البيانات الكبيرة تحت مظلة مبادرة "عربستات"، تكون كمنصة إقليمية يلعب فيها صندوق النقد العربي دوراً قيادياً في مجال البيانات الكبيرة، وتهدف لتبادل التجارب والخبرات مما سيساهم في تعزيز الاستفادة من هذه البيانات في الإحصاءات الرسمية ولأغراض الإقتصاد الكلي وكذلك وضع السياسات المناسبة. يمكن أن تشمل:

- عقد ورش عمل ومؤتمرات إقليمية
- إنشاء بوابة إلكترونية مختصة بقضايا البيانات الكبيرة.
- تبادل تجارب الدول في مجال البيانات الكبيرة.
- بناء شراكات وتعاون على المستوى الإقليمي والمستوى الدولي.

ستساعد الشبكة على مناقشة مجالات التعاون بين الأطراف المعنية والاستفادة من الخبرات الفنية الموجودة في المراكز الإحصائية والبنوك المركزية ووزارات المالية في الدول العربية، ومطابقة واستعراض

استراتيجية الموارد البشرية المتخصصة في مجال البيانات الكبيرة. كذلك التركيز على خلق روابط تعاون وشركات مع منظمات إقليمية ودولية ومزودي المعلومات في مجال البيانات الكبيرة، وتوفير مجال لتبادل المعلومات والخبرات، وكذلك العمل على تحقيق أكبر تطبيقات للبيانات الكبيرة بأقل التكاليف علاوة على توفير الوقت المستهلك للتغلب على التحديات في هذا المجال.

شكل 34: إنشاء شبكة إقليمية للتعاون وبناء الشراكات في مجال البيانات الكبيرة تحت مظلة مبادرة "عربستات"

7. الخلاصة والتوصيات

خلصت دراسة تطبيقات البيانات الكبيرة في الدول العربية، التي شملت الجهات المعنية في وزارات المالية والمصارف المركزية ومؤسسات النقد العربية وأجهزة الإحصاء العربية، إلى تقييم الوضع الراهن وجاهزية الجهات العربية في مجال البيانات الكبيرة، ورصد احتياجاتها لإنجاز وتطوير مشاريع البيانات الكبيرة، كذلك جمع معلومات واسعة حول مصادر البيانات وأدوات التحليل التي تستخدمها الجهات، ومشاريعها الحالية والمستقبلية في مجال البيانات الكبيرة، والشراكات مع مزودي البيانات ومع المنظمات الدولية.

بينت نتائج استبيان تطبيقات البيانات الكبيرة في الدول العربية اهتمام الجهات المعنية بموضوع البيانات الكبيرة ودعم تحليلات الاقتصاد الكلي. على الرغم من المشاركة الفعلية لبعض الجهات في تطبيقات البيانات الكبيرة، إلا أن جلها ترغب في التعاون مع نظرائها، والمشاركة مع دول أو مؤسسات أخرى أو لجان أو ورش عمل في مجال البيانات الكبيرة، كذلك المساهمة في وضع خارطة طريق و/أو خطة للتعاون الإقليمي تحت مظلة مبادرة "عربستات" حول استخدام وتطوير البيانات الكبيرة وإدارة التحديات المرتبطة بها، مما سيساعدها في عملية استكشاف البيانات الكبيرة وتحليلها، الأمر الذي تعتبره أغلبية الجهات مهمة معقدة.

حددت الجهات المستطلعة عدداً من المجالات ذات أولوية للتركيز عليها، تغطي مجموعة واسعة من المواضيع الإحصائية البحتة، مثل جودة البيانات، وتقنيات أخذ العينات، وتطوير الإحصاءات والمؤشرات،

وغيرها. كذلك، أكدت الجهات عن التحديات التي تواجه استخدام البيانات الكبيرة أو البدء في مشاريع البيانات الكبيرة، على رأسها إمكانية الوصول إلى البيانات الكبيرة والحصول عليها بصفة دورية، والحاجة إلى المهارات التقنية والخبرات الوطنية في مجال البيانات الكبيرة، كذلك قضايا التشريعات المتعلقة بسرية وخصوصية البيانات، إضافة إلى التعاون مع الجهات المنتجة للبيانات وإبرام مذكرات تفاهم حول تبادل البيانات، إلى جانب الإمكانات المادية لتنفيذ مشروعات البيانات الكبيرة واستمراريتها.

على ضوء نتائج استبيان تطبيقات البيانات الكبيرة والتحديات التي تواجه تطوير الجهات المُستطلعة لمشاريع وشراكات في هذا المجال، تم اقتراح وضع إطار شامل لبرنامج عمل إقليمي لدعم تطبيق البيانات الكبيرة في الدول العربية تحت مظلة مبادرة "عربستات"، يتضمن ثلاث محاور رئيسية، وهي: إنشاء فريق عمل إقليمي معني بالبيانات الكبيرة، وإطلاق مبادرة إقليمية لبناء القدرات في مجال البيانات الكبيرة، كذلك إنشاء شبكة إقليمية للتعاون وتبادل الخبرات وبناء الشراكات في مجال البيانات الكبيرة.

خلصت الدراسة إلى توصية جعل استراتيجية البيانات الكبيرة جزء من استراتيجية الحكومات على نطاق أوسع واعتماد خطة عمل وخارطة طريق تشمل بناء القدرات، ودراسة إمكانيات مصادر البيانات الكبيرة، وتطوير المهارات، وقضايا الخصوصية وحماية البيانات، والتعاون على المستوى العربي والعالمي، والشراكات بين مختلف أصحاب المصلحة (الحكومة والأكاديميين والقطاع الخاص)، وتتبع التطورات في المنهجية، وتقييم الجودة وتكنولوجيا المعلومات.

التوصيات

التوصية الأولى. تعزيز استراتيجية البيانات الكبيرة كجزء من استراتيجية الحكومات على نطاق أوسع.

التوصية الثانية. العمل على تنفيذ الإطار المقترح لبرنامج عمل إقليمي لدعم تطبيق البيانات الكبيرة في الدول العربية تحت مظلة مبادرة "عربستات".

التوصية الثالثة. إنشاء فريق عمل إقليمي لتعزيز تطبيقات البيانات الكبيرة في الدول العربية تحت مظلة مبادرة "عربستات".

التوصية الرابعة. إطلاق مبادرة إقليمية لبناء القدرات في مجال البيانات الكبيرة في معهد السياسات الاقتصادية التابع لصندوق النقد العربي، وتقديم المشورة بشأن برنامج تدريب يعالج الثغرات والاحتياجات للمهارات في الدول الأعضاء ضمن مهام فريق العمل الإقليمي المعني بالبيانات الكبيرة.

التوصية الخامسة. إنشاء شبكة إقليمية للتعاون وبناء الشراكات في مجال البيانات الكبيرة تحت مظلة مبادرة "عربستات".

التوصية السادسة. التعاون مع جهات إقليمية ودولية وخاصة مع خبراء فريق عمل الأمم المتحدة المعني بالبيانات الكبيرة، استناداً لنتائج استبيان تطبيقات البيانات الكبيرة في الدول العربية، وتقييم مدى استعداد الجهات المعنية في المنطقة العربية واحتياجاتها لاستخدام البيانات الكبيرة. كذلك المساهمة في أنشطة فريق العمل الإقليمي المقترح بشأن برامج وفعاليات البيانات الكبيرة، وتنسيق خطط التدريب مع فريق العمل الإقليمي المقترح بما يتماشى مع الدراسة الاستقصائية لصندوق النقد العربي حول تطبيقات البيانات الكبيرة في الدول العربية، إلى جانب المشاركة في الاجتماعات السنوية لفريق العمل تحت مظلة مبادرة "عربستات".

التوصية السابعة: إنشاء بوابة إلكترونية تختص بقضايا البيانات الكبيرة على موقع صندوق النقد العربي.

التوصية الثامنة: دعم الاستراتيجية في المجالات الإحصائية الناشئة، واستخدام مصادر البيانات الإدارية لأغراض إحصائية في الدول العربية، كذلك تطوير مؤشرات العلم والتقنية (science and technology)، وإدماج المعلومات الإحصائية والجغرافية المكانية.

التوصية التاسعة: تفعيل عضوية الدول العربية وصندوق النقد العربي في اللجان الدولية المعنية بالبيانات الكبيرة.

8. المرفقات

1.8 استبيان تطبيقات البيانات الكبيرة في الدول العربية

صندوق النقد العربي

ARAB MONETARY FUND

صندوق النقد العربي

استبيان حول

"استخدام البيانات الكبيرة في الدول العربية"

(Big Data Questionnaire)

استبيان حول

"استخدام البيانات الكبيرة في الدول العربية"

أولاً: بيانات عامة حول المؤسسة

الدولة:

المؤسسة:

الدائرة أو القسم أو العامل على البيانات الكبيرة:

نقطة الاتصال:
الاسم:
المسمى الوظيفي:
الهاتف:
البريد الإلكتروني الرسمي:

ثانياً: استراتيجية البيانات الكبيرة والخبرة العامة

1. **هل قامت مؤسستكم بإعداد خطة أو استراتيجية لاستخدام البيانات الكبيرة؟**
   - ☐ **نعم**
   - ☐ **لا**

2. **هل تم تشكيل لجنة او وحدة فنية متخصصة بموضوع البيانات الكبيرة في مؤسستكم؟**
   - ☐ **نعم، (لجنة داخلية)**
   - ☐ **نعم، (لجنة بالتنسيق مع الشركاء)**
   - ☐ **لا**

3. **هل لدى مؤسستكم عضوية في لجان إقليمية أو دولية في مجال البيانات الكبيرة؟**
   - ☐ **نعم**

☐ لا

4. ما هو مستوى اهتمام الادارة العليا في مؤسستكم بموضوع البيانات الكبيرة؟

☐ 1 (اهتمام ضعيف)

☐ 2

☐ 3

☐ 4

☐ 5 (اهتمام كبير)

5. هل تعملون حالياً على أي مشروع للبيانات الكبيرة؟

☐ نعم

☐ لا

إذا كان الجواب نعم، يرجى تحديده في الجزء الثالث من الاستبيان. (سؤال رقم 29)

---

6. هل تنظر مؤسستكم لبدء مشاريع لها صلة بالبيانات الكبيرة في الفترة المقبلة؟

☐ نعم

☐ لا

7. إذا كانت الإجابة نعم، ما الغرض من هذه المشاريع؟

---

8. هل تستخدم مؤسستكم مصادر البيانات الكبيرة؟

☐ نعم

☐ لا

في حال الإجابة "لا"، الرجاء تحديد الدوافع حول عدم استخدام أي مصادر للبيانات الكبيرة

---

9. ما مصادر البيانات الكبيرة التي تستخدمها مؤسستكم؟

---

10. هل لدى مؤسستكم إمكانية الوصول المباشر إلى مصادر البيانات الكبيرة؟

- ☐ نعم
- ☐ لا

11. استخدام البيانات الإدارية:

أ. هل تستخدمون البيانات الإدارية الرسمية؟ (البيانات الإدارية هي: البيانات التي توفرها السجلات الإدارية في المؤسسات الحكومية)

- ☐ نعم
- ☐ لا

ب. إذا كان الجواب بنعم، ماهي مصادر البيانات الادارية المعتمدة لديكم؟

ت. ما كيفية الحصول عليها؟

- ☐ مباشرة من المصدر
- ☐ عن طريق جهاز الاحصاء
- ☐ طرق أخرى، يرجى تحديدها ----------

12. ما هي برأيكم أكبر خمسة تحديات تواجه مؤسستكم في استخدام البيانات الكبيرة أو البدء في مشاريع البيانات الكبيرة؟

- ☐ 1 (أكبر تحدي)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 (أقل تحدي)

13. ما تقييمكم للجودة الإحصائية لمصادر البيانات الكبيرة ومخرجات التحليل المستندة إلى مصادر البيانات الكبيرة؟

مصادر البيانات الكبيرة	مخرجات التحليل
☐ 1 (ضعيف)	☐ 1 (ضعيف)
☐ 2	☐ 2
☐ 3	☐ 3
☐ 4	☐ 4
☐ 5 (عالي جداً)	☐ 5 (عالي جداً)

14. ما المخرجات التي تتوقعونها نتيجة استخدام البيانات الكبيرة (على سبيل المثال القدرة على التنبؤ الآني) - إن توفرت؟

---

15. ما القضايا المرتبطة بالسرية والخصوصية التي تواجهها مؤسستكم فيما يتعلق بمصادر البيانات الكبيرة؟

---

16. هل لديكم إطار قانوني لمعالجة قضايا السرية والخصوصية في البيانات الكبيرة؟

☐ نعم

☐ لا

17. ما المهارات التي تعتبرونها الأهم بالنسبة للموظفين الذين يتعاملون مع البيانات الكبيرة؟

---

18. كيف عالجتم و / أو تخططون لمعالجة النقص في المهارات في البيانات الكبيرة لدى مؤسستكم؟

---

19. هل تقومون بتوظيف خبراء متخصصين في البيانات؟ هل تقومون بتدريب الموظفين الحاليين في مجال البيانات الكبيرة؟ ما الخيار الأكثر فائدة لكم؟

---

20. هل لديكم اتفاقيات تعاون/ إجراءات عمل لشراكات مع مزودي البيانات؟

- ☐ نعم
- ☐ لا

21. هل مؤسستكم على استعداد للمشاركة مع دول او مؤسسات أخرى أو لجان أو ورش عمل حول موضوع البيانات الكبيرة؟

- ☐ نعم
- ☐ لا

إذا كانت الإجابة نعم، يرجى ذكر مجالات مشاركتكم مع الجهات الأخرى؟

---

22. هل مؤسستكم على استعداد للمساهمة في وضع خارطة طريق/ أو خطة للتعاون الإقليمي حول استخدام وتطوير البيانات الكبيرة؟

- ☐ نعم
- ☐ لا

23. ما المواضيع الإحصائية التي تهمكم في إطار البيانات الكبيرة؟

يرجى اختيار كل ما ينطبق:

- ☐ التعرف على مصادر البيانات الكبيرة
- ☐ تقنيات أخذ العينات
- ☐ جودة البيانات
- ☐ الإحصاءات والمؤشرات
- ☐ استخدام التقنيات الإحصائية لإدارة البيانات الكبيرة
- ☐ طرق التجميع
- ☐ التنبؤات والأداء الاقتصادي
- ☐ تحليل السيناريوهات والافتراضات
- ☐ تعزيز المنافسة
- ☐ استخدام البيانات لاتخاذ القرار
- ☐ تحديات التواصل

- ☐ التكاليف والمهارات والاحتياجات من الموارد
- ☐ أخرى (يرجى تحديدها) ----------

24. كيف تقيمون مدى استعداد مؤسستكم لبدء تحليل البيانات الكبيرة أو الحصول على المخرجات المتوقعة من استخدامها بشكل منتظم؟

- ☐ 1 (ضعيف)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 (عالي)

25. ما الأدوات التي تستخدمونها لتحليل البيانات الكبيرة؟

---

26. كيف يتم تطوير هذه الأدوات؟

- ☐ داخلياً
- ☐ عن طريق الاستعانة بمصادر خارجية
- ☐ أو الاثنين معاً

27. يرجى وصف بعض الآثار المحتملة للموارد على مؤسستكم في بدء تحليل البيانات الكبيرة و / أو الحصول على المخرجات المتوقعة من استخدامها بشكل منتظم، لا سيما من حيث موارد تقنية المعلومات (كل من البرمجيات والأجهزة) ورأس المال البشري (فيما يخص تقنية المعلومات، والإحصاءات، والبحوث الاقتصادية، وعمليات التصنيف، وأمن البيانات، والأمن الوطني).

---

28. هل لديكم أية مقترحات أخرى تودون ذكرها لتطوير البيانات الكبيرة، من تشريعات وموارد بشرية وتقنية المعلومات، والوصول إلى البيانات وما إلى ذلك؟

ثالثاً: مشاريع البيانات الكبيرة - لمحة موجزة عن مشاريع البيانات الكبيرة التي تعمل مؤسستكم على إنجازها

29. اسم او عنوان المشروع: عنوان وصفي أو فكرة رئيسية للمشروع أو المشروع المقترح

---

30. ما الوضع الراهن للمشروع؟

- ☐ مرحلة الفكرة
- ☐ مقترح (في التخطيط - لم تتم الموافقة عليه أو تمويله بعد)
- ☐ تمت الموافقة عليه (تمت الموافقة عليه - لم يتم تمويله بعد)
- ☐ ممول (معتمد وممول – لم يبدأ بعد)
- ☐ جاري (في مرحلة التنفيذ)
- ☐ اكتمل

31. مجالات الاستخدام المحتملة لهذا المشروع: حدد كل ما ينطبق.

- ☐ مرحلة الفكرة
- ☐ الإحصاءات الديمغرافية والاجتماعية
- ☐ الإحصاءات الاقتصادية والمالية
- ☐ الإحصاءات البيئية
- ☐ إحصاءات تكنولوجيا المعلومات والاتصالات
- ☐ إحصاءات العمل
- ☐ إحصاءات التعليم
- ☐ إحصاءات الصحة
- ☐ إحصاءات الأسعار
- ☐ إحصاءات السياحة
- ☐ إحصاءات النقل
- ☐ مجالات أخرى للإحصاءات الرسمية: يرجى التحديد ----------

32. المشروع مؤهل على النحو التالي:

- ☐ استكشافي / للبحث
- ☐ مشروع "Pilot" بهدف تطبيقه في حال نجاح المشروع
- ☐ لإعداد وإصدار الإحصاءات
- ☐ أخرى (يرجى التحديد)

33. الشراكات:

أ. هل لدى مؤسستكم أي شراكات مع منظمات أو موفري بيانات آخرين حول هذا المشروع؟

- ☐ نعم
- ☐ لا

ب. اسم الشريك: في حال عدم الرغبة في الكشف عن الإسم، يرجى تقديم تصنيف عملي - على سبيل المثال. "شريك - مزود بيانات الهاتف المحمول".

ج. نوع المؤسسة الشريكة: حدد كل ما ينطبق.

- ☐ مؤسسة دولية حكومية
- ☐ مؤسسة دولية غير حكومية
- ☐ مؤسسة تجارية غير حكومية
- ☐ مؤسسة أكاديمية
- ☐ أخرى (يرجى التحديد) ----------

د. نوع الشراكة: حدد كل ما ينطبق

- ☐ مزود البيانات
- ☐ شريك التكنولوجيا لتصميم و هيكلة البيانات
- ☐ الشريك التحليلي
- ☐ غير ذلك (يرجى التحديد)

ه. الوضع الحالي للشراكة:

- ☐ في المناقشة
- ☐ النماذج الأولية / الاختبار (يسمح بعض شركاء البيانات بذلك قبل توقيع العقد)
- ☐ عقد اتفاقية

☐ غير ذلك (يرجى التحديد) ----------

34. نأمل إضافة ما ترون مناسبته حيال الموضوع:

ملاحظـــة

الرجاء التكرم بإعادة الاستبيان بعد تعبئته إلى:

الدائرة الاقتصادية والفنية
صندوق النقد العربي – أبوظبي
بريد الكتروني: FinancialMarkets@amfad.org.ae
في حال وجود سؤال الرجاء التواصل معنا على البريد الالكتروني soumia.lotfi@amf.org.ae
هاتف: 6171552 / 97126171556+

2.8 المراجع

1. The International Monetary Fund. Big Data: Potential, Challenges, and Statistical Implications 2017.
2. Charles Bean. Independent Review of UK Economic Statistics 2016.
3. Irem Önder. Forecasting tourism demand with Google trends: Accuracy comparison of countries versus cities, International Journal of Tourism Research, 2017.
4. Paul Smith. Google's MIDAS Touch: Predicting UK Unemployment with Internet Search Data, Journal of Forecasting, 2016.
5. Alessandro Girardi, Christian Gayer, and Andreas Reut. The Role of Survey Data in Nowcasting Euro Area GDP Growth, Journal of Forcasting, 2015.
6. Hal Varian & Choi.Predicting the Present with Google Trends, 2009.
7. Hal Varian. Big Data: New Tricks for Econometrics, 2013.
8. Zimmermann K & Askitas N. Google Econometrics and Unemployment Forecasting, 2009.
9. D'Amuri F, & Marcucci J. "Google it!"Forecasting the US Unemployment Rate with a Google Job Search index, 2010. http://www.feem.it/userfiles/attach/20103261238204NDL2010-031.pdf
10. D'Amuri F, & Marcucci J. The predictive power of Google searches in forecasting unemployment, 2013. http://www.bancaditalia.it/pubblicazioni/temi-disc ... 0891/en_tema_891.pdf
11. McLaren N & Shanbhogue R. Using internet search data as economic indicators, 2011. http://www.bankofengland.co.uk/publications/quarterlybulletin/qb110206.pdf
12. Simeon Vosen and Torsten Schmidt. Forecasting private consumption: survey-based indicators vs. Google trends, 2011. http://hdl.handle.net/10.1002/for.1213
13. Yan Carrière-Swallow and Felipe Labbé. Nowcasting with google trends in an emerging market. Journal of Forecasting, 2011. doi: 10.1002/for.1252. URL http://ideas.repec. org/p/chb/bcchwp/588.html.Working Papers Central Bank of Chile 588.
14. Fantazzini D & Toktamysova Z (2014). Forecasting German Car Sales Using Google Data and Multivariate Models, 2014. https://mpra.ub.uni-muenchen.de/67110/1/MPRA_paper_67110.pdf
15. A. Cavallo and R. Rigobon: "The Billion Prices Project: Using Online Prices for Measurement and Research", Journal of Economic Perspectives, volume 30(2), spring 2016, p. 151.
16. J. de Haan, L. Willenborg and A.G. Chessa: "An overview of price index methods for scanner data", room document for the meeting of the UNECE Group of Experts on Consumer Price Indices, May 2016.
17. The Australia's welfare indicator framework. Australian Government. Australian Institute of Health and Welfare, 2017.
18. Ekaterina Olshannikova, Thomas Olsson, Jukka Huhtamäki and Hannu Kärkkäinen. Conceptualizing Big Social Data, Journal of Big Data, 2017.
19. Gandomi A, Haider M. Beyond the hype: big data concepts, methods, and analytics. Int J Inf Manag. 2015;35(2):137–44.
20. Aradau C, Blanke T. Politics of prediction: security and the time/space of governmentality in the age of big data. European Journal of Social Theory. 2016:1–19. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/1368431016667623.
21. Saldana-Perez AMM, Moreno-Ibarra M. Traffic analysis based on short texts from social media. Int J Knowl Soc Res. 2016;7(1):63–79.

22. Bravo-Marquez F, Mendoza M, Poblete B. Meta-level sentiment models for big social data analysis. Knowl Based Syst. 2014;69:86–99.
23. Pandarachalil R, Sendhilkumar S, Mahalakshmi G. Twitter sentiment analysis for large-scale data: an unsupervised approach. Cogn Comput. 2015;7(2):254–62.
24. Ishikawa H. Social big data mining. Boca Raton: Taylor & Francis Group, CRC Press; 2015.
25. Chen M, Mao S, Zhang Y, Leung VC. Big data: related technologies, challenges and future prospects. In: Springerbriefs in computer science. Cham: Springer; 2014.
26. Moorthy J, Lahiri R, Biswas N, Sanyal D, Ranjan J, Nanath K, Ghosh P. Big data: prospects and challenges. J Decis Makers. 2015;40:74–96.
27. Guellil I, Boukhalfa K. Social big data mining: a survey focused on opinion mining and sentiments analysis. In: 2015 12th international symposium on programming and systems (ISPS). New York: IEEE; 2015. p. 1–10.
28. Duan L, Xiong Y. Big data analytics and business analytics. J Manag Anal. 2015;2(1):1–21.

3.8 مرفق - الأشكال والجداول

قائمة الأشكال

شكل 1. نسبة المشاركين في الاستبيان 13
شكل 2. هل قامت مؤسستكم بإعداد خطة أو استراتيجية لاستخدام البيانات الكبيرة؟ 13
شكل 3. هل تم تشكيل لجنة أو وحدة فنية متخصصة بموضوع البيانات الكبيرة في مؤسستكم؟ 14
شكل 4. هل لدى مؤسستكم عضوية في لجان إقليمية أو دولية في مجال البيانات الكبيرة؟ 14
شكل 5. مستوى إهتمام الإدارة العليا في مؤسستكم بموضوع البيانات الكبيرة 14
شكل 6. هل تعملون حالياً على أي مشروع للبيانات الكبيرة؟ 14
شكل 7. هل تنظر مؤسستكم لبدء مشاريع لها صلة بالبيانات الكبيرة في الفترة المقبلة 15
شكل 8. هل تستخدم مؤسستكم مصادر البيانات الكبيرة؟ 16
شكل 9. مصادر البيانات الكبيرة التي تستخدمها مؤسستكم 17
شكل 10. هل لدى مؤسستكم إمكانية الوصول المباشر إلى مصادر البيانات الكبيرة؟ 17
شكل 11. هل تستخدمون البيانات الإدارية الرسمية؟ 17
شكل 12. مصادر البيانات الادارية المعتمدة 18
شكل 13. كيف تحصل مؤسستكم على البيانات الإدارية؟ 18
شكل 14. ما هو تقييمكم للجودة الإحصائية لمصادر البيانات الكبيرة؟ 20
شكل 15. ما هو تقييمكم للجودة الإحصائية لمخرجات تحليل البيانات الكبيرة؟ 20
شكل 16. هل لديكم إطار قانوني لمعالجة قضايا السرية والخصوصية في البيانات الكبيرة؟ 22
شكل 17. هل تقومون بتوظيف خبراء متخصصين في مجال البيانات الكبيرة؟ 23
شكل 18. هل لديكم اتفاقيات تعاون/ إجراءات عمل لشراكات مع مزودي البيانات الكبيرة؟ 23
شكل 19. المواضيع الإحصائية التي تهم الجهات المعنية في إطار البيانات الكبيرة 25
شكل 20. تقييم مدى استعداد مؤسستكم لبدء تحليل البيانات الكبيرة أو الحصول على المخرجات المتوقعة من استخدامها بشكل منتظم؟ 25
شكل 21. التوفر على أدوات لتحليل البيانات الكبيرة 25
شكل 22. كيف يتم تطوير الأدوات المستخدمة لتحليل البيانات الكبيرة؟ 26
شكل 23. مقترحات الجهات المستطلعة في الدول الأعضاء لتطوير تطبيقات البيانات الكبيرة من تشريعات، وموارد بشرية، وتقنية المعلومات، والوصول إلى البيانات وغيرها 29
شكل 24. الوضع الراهن لمشاريع البيانات الكبيرة 31
شكل 25. مجالات الاستخدام المحتملة لمشاريع البيانات الكبيرة 31
شكل 26. مشاريع البيانات الكبيرة المؤهلة 31
شكل 27. وجود شراكات مع منظمات أو موفري بيانات آخرين حول مشروع البيانات الكبيرة 31

شكل 28. نوع المؤسسات الشريكة 32
شكل 29. نوعية الشراكات 32
شكل 30. الوضع الحالي للشراكات 32
شكل 31. بحيرة أو مركز البيانات "Data Lake or Data Hub" 40
شكل 32. الإطار العام لبرنامج عمل إقليمي لدعم تطبيق البيانات الكبيرة في الدول العربية 42
شكل 33. مبادرة إقليمية لبناء القدرات في مجال البيانات الكبيرة 44
شكل 34. إنشاء شبكة إقليمية للتعاون وبناء الشراكات في مجال البيانات الكبيرة 46

قائمة الجداول

جدول 1. قياسات حجم البيانات 9
جدول 2. بعض مصادر البيانات الكبيرة 9
جدول 3. أمثلة من الواقع حول النمو السريع للبيانات الكبيرة/البيانات غير المهيكلة 10
جدول 4: أكبر خمسة تحديات تواجه الجهات العربية المعنية في استخدام البيانات الكبيرة أو البدء في مشاريع البيانات الكبيرة 19
جدول 5: الأدوات التي تستخدمها الجهات المستطلعة لتحليل البيانات الكبيرة 26
جدول 6. مشاريع البيانات الكبيرة - الاسم والوصف 30
جدول 7: أمثلة عن الدراسات التي تطرقت لجوانب الاقتصاد الكلي باستخدام البيانات الكبيرة 33
جدول 8: الإنجازات والتقدم المحرز لفريق العمل العالمي للأمم المتحدة المعني بالبيانات الكبيرة 34
جدول 9: بعض الأدوات المستخدمة في مجال البيانات الكبيرة 38