صندوق النقد العربي
ARAB MONETARY FUND

تطبيقات البيانات الكبيرة "Big Data"
في الدول العربية

د. سومية لطفي

أمانة اللجنة الفنية لمبادرة عربستات

نوفمبر 2017

نقاط العرض

1. هدف الدراسة ونطاقها
2. مفهوم البيانات الكبيرة والفرص والتحديات
3. غرض ونتائج استبيان تطبيقات البيانات الكبيرة في الدول العربية
4. بعض تطبيقات البيانات الكبيرة لأغراض الإحصاءات الرسمية وتحليل الاقتصاد الكلي
5. مقترحات لوضع إطار لبرنامج عمل إقليمي لدعم تطبيق البيانات الكبيرة في الدول العربية

1. هدف الدراسة ونطاقها

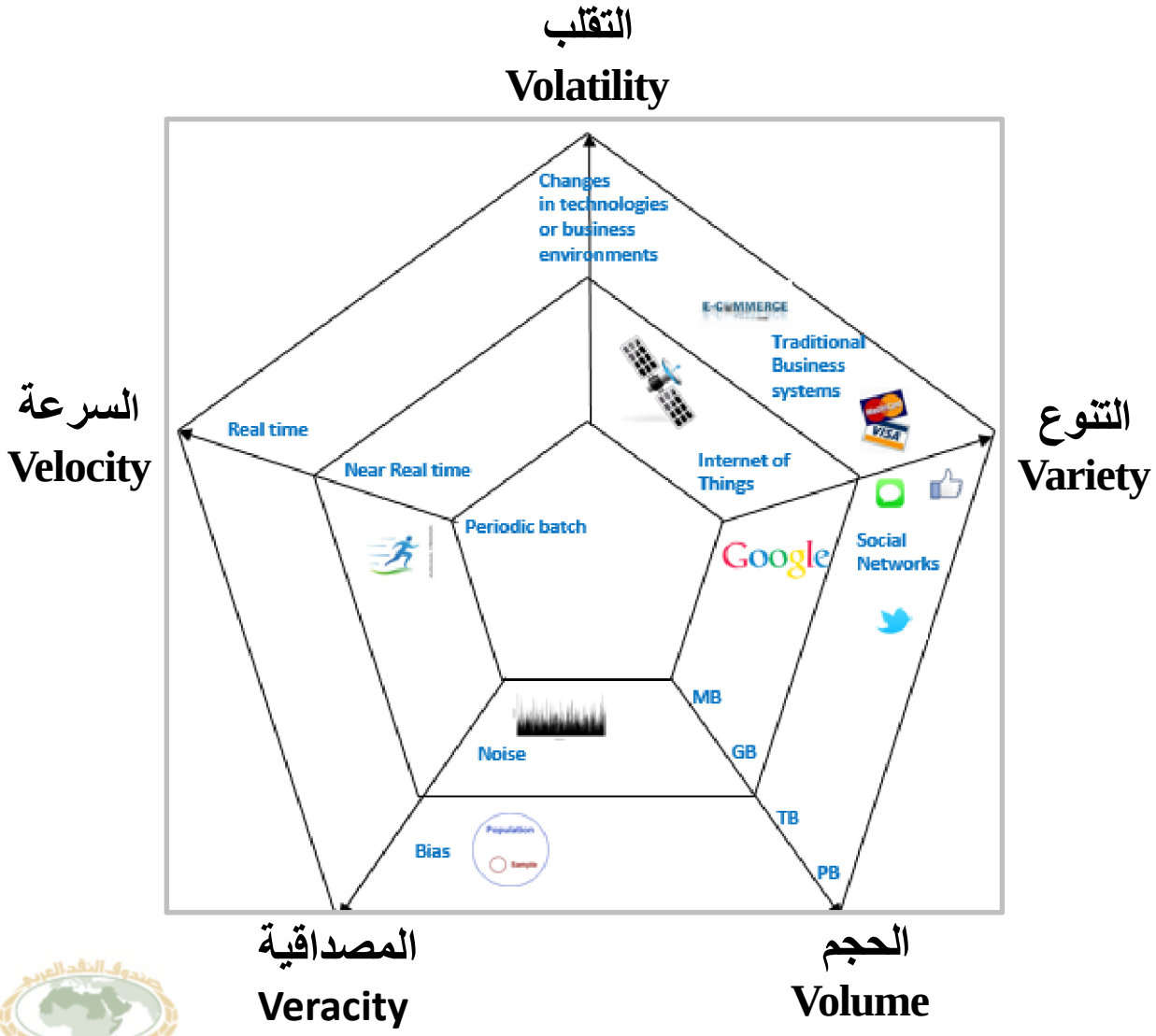
تقييم الوضع الراهن وجاهزية السلطات العربية في مجال البيانات الكبيرة استناداً إلى استبيان تطبيقات البيانات الكبيرة الذي شمل الجهات المعنية في وزارات المالية، والمصارف المركزية ومؤسسات النقد العربية، وأجهزة الإحصاء في الدول العربية.

استعراض بعض الممارسات بشأن البيانات الكبيرة لأغراض الإحصاءات الرسمية و تحليل الاقتصاد الكلي.

تقديم مقترح لوضع إطار لبرنامج عمل إقليمي لدعم تطبيق البيانات الكبيرة في الدول العربية.

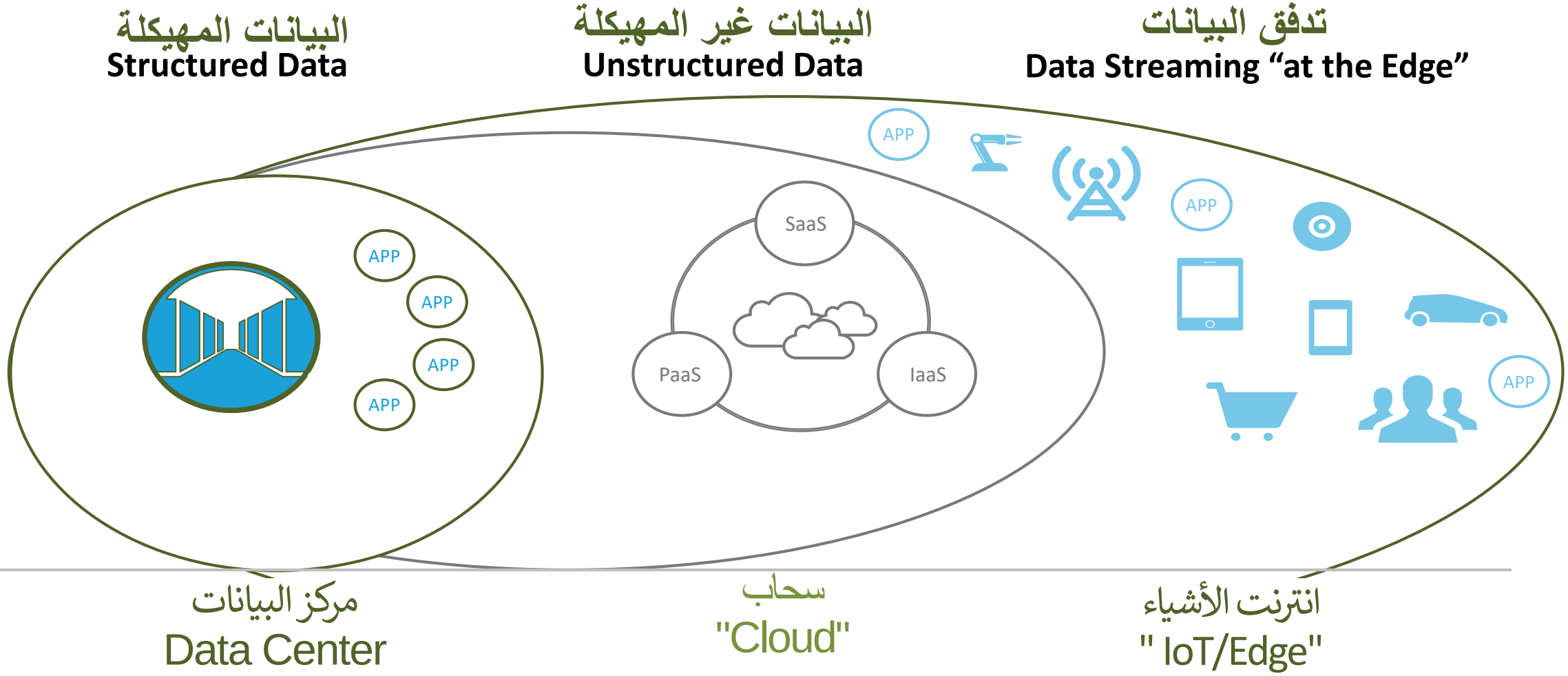
تعريف البيانات الكبيرة - مبدء "5 V's"

- التعريف الواسع النطاق للبيانات الكبيرة: بيانات ذات حجم كبير، وتشمل مجموعة متنوعة من البيانات، الأمر الذي يتطلب أدوات وطرق جديدة لتجميعها، وإدارتها، ومعالجتها بطريقة فعالة.
- تعريف البيانات الكبيرة وفقاً لمبدئ "5 V's": الحجم - التنوع - السرعة - المصداقية - التقلب.



المصدر: صندوق النقد الدولي

البيانات الكبيرة - منظور واسع



المصدر: "Cisco“

البيانات الكبيرة والإحصاءات الرسمية

مدى تأثر الإحصاءات الرسمية بجمع البيانات المتدفقة من خلال شبكات

- الاتصالات المتنقلة وأجهزة الأقمار الصناعية، والصور الإلكترونية.
- "Internet Of Things"
- وسائل الاعلام الاجتماعية: الأحداث والآراء التي يتم تجميعها من محركات البحث، مثل وسائل الاعلام الاجتماعية: " Twitter, blogs, text messages “Facebook, LinkedIn
- الأجهزة القابلة للارتداء
- الانسان الآلي "Robot" ومستشعر المعلومات "Sensor" في عمليات الإنتاج
- حركة المرور الذاتية
- الأنظمة الذكية
- مدفوعات الائتمان والخصم، منصات التداول والتسوية
- رسوم الطرق، حركة المسافرين ومسارات القطارات والسفن والطائرات، أجهزة تتبع المحمول، أنظمة الملاحة
- الكلام والتعرف على الكلمة

...

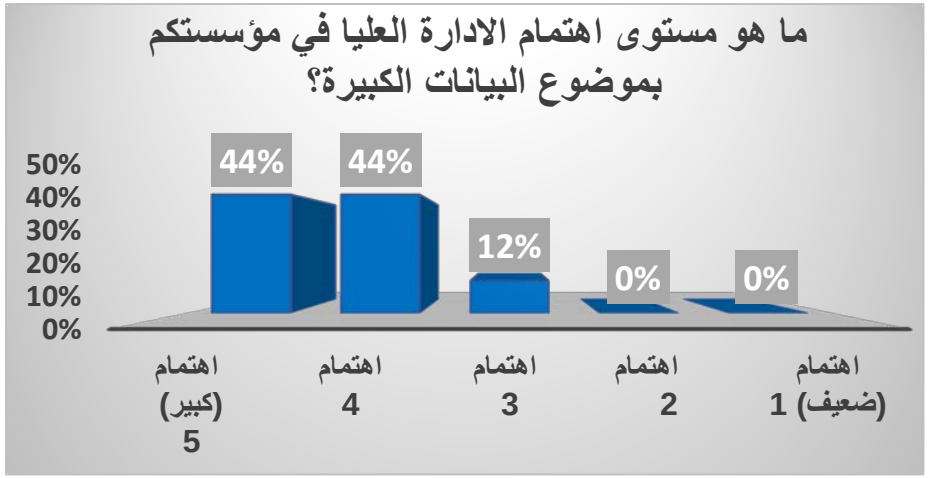
3. غرض ونتائج استبيان تطبيقات البيانات الكبيرة في الدول العربية

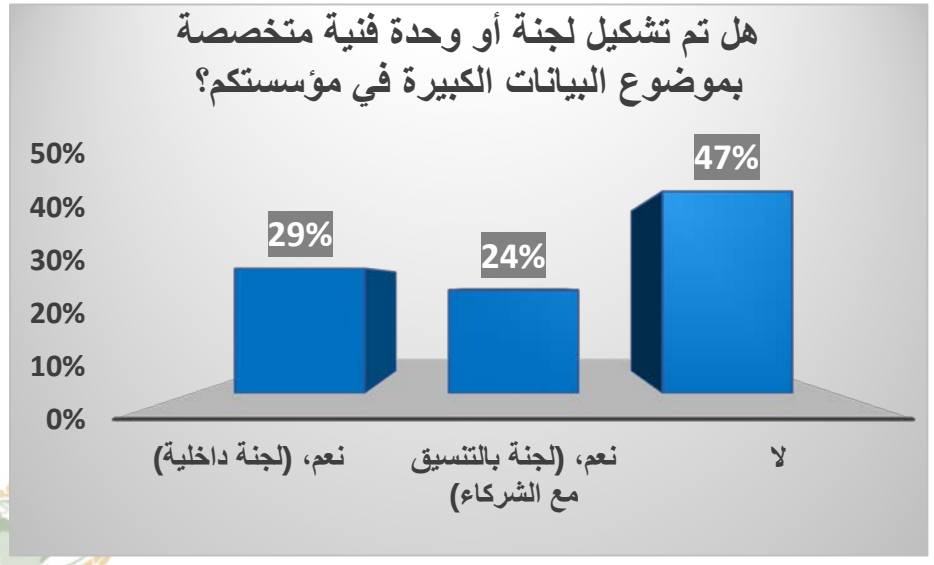
صندوق النقد العربي
ARAB MONETARY FUND

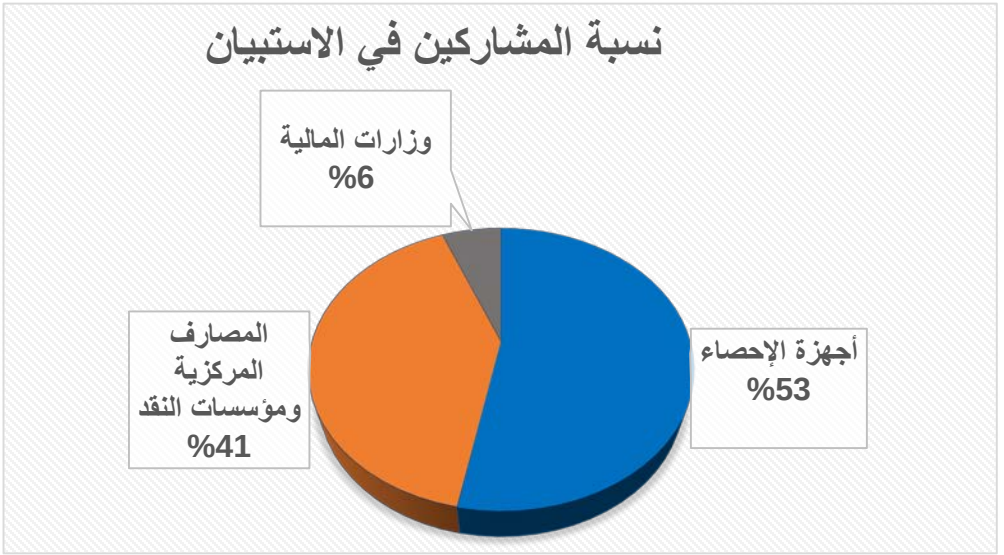
الغرض من استبيان تطبيقات البيانات الكبيرة في الدول العربية

اهتمام كبير من قبل الجهات المعنية في الدول العربية بموضوع البيانات الكبيرة

- تم استلام 17 ردَ حول الاستبيان من قبل ممثلي الدوائر المعنية في وزارات المالية، والمصارف المركزية ومؤسسات النقد العربية، والمراكز الإحصائية في الدول العربية.
- حوالي 88% من الردود المستلمة أشارت إلى اهتمام الادارة العليا بموضوع البيانات الكبيرة، وحوالي 29% منها قام بتكوين لجنة داخلية تناقش وتستعرض هذا المجال.

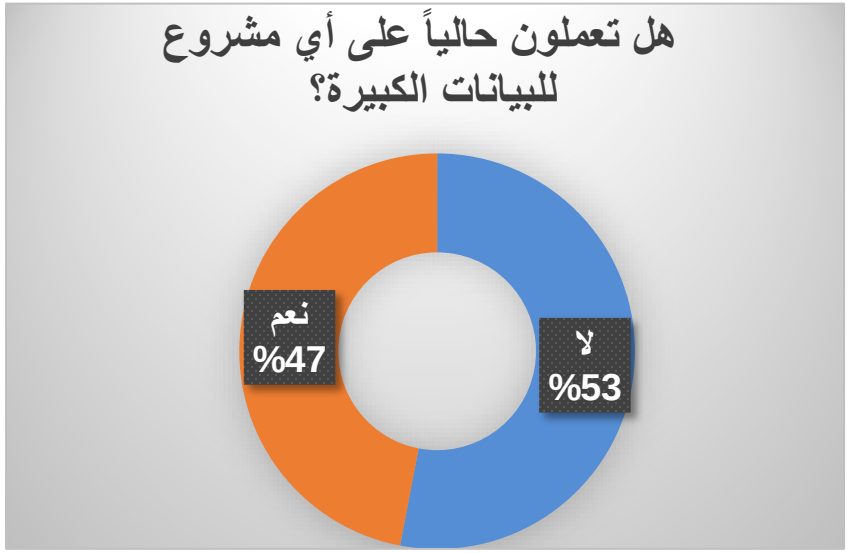


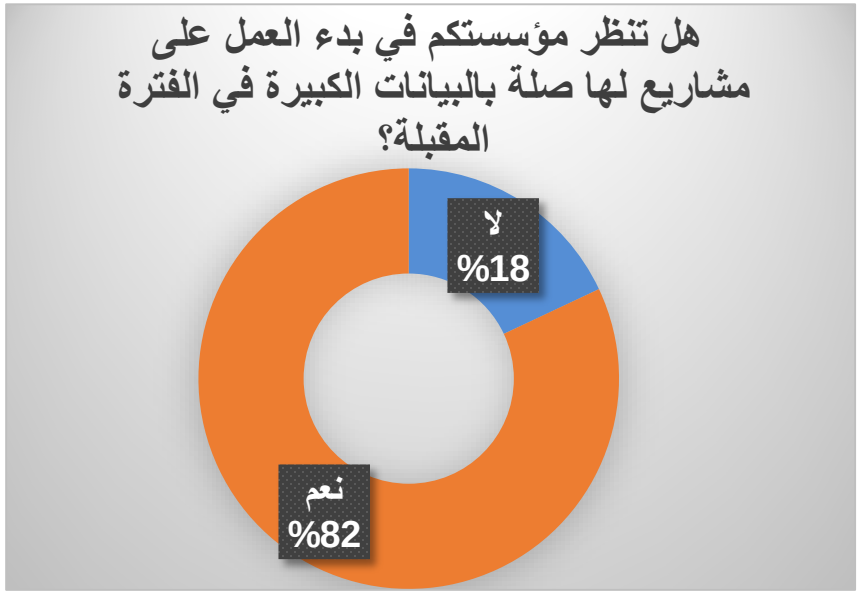




المشاركة الفعلية للجهات المعنية في مجال البيانات الكبيرة

- يعمل حوالي 47% من المؤسسات المعنية العربية على مشاريع البيانات الكبيرة حالياً (34% منها تمت الموافقة عليها، لكن لم يتم تمويلها بعد).



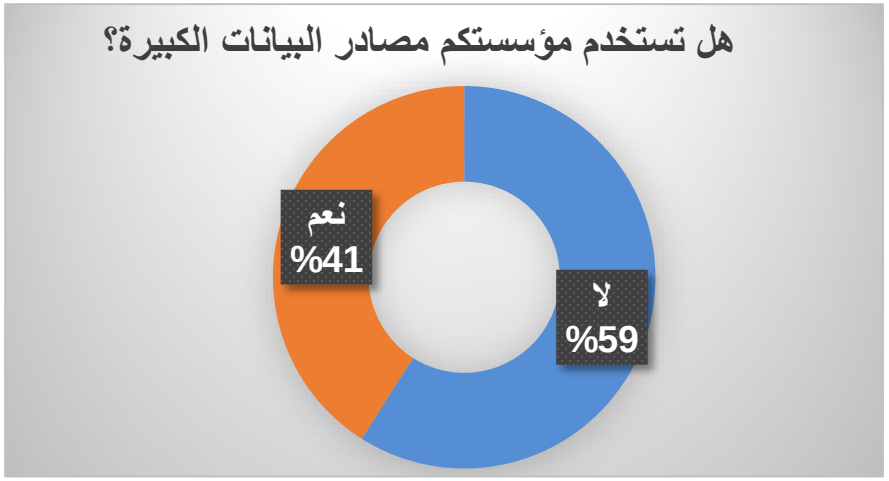


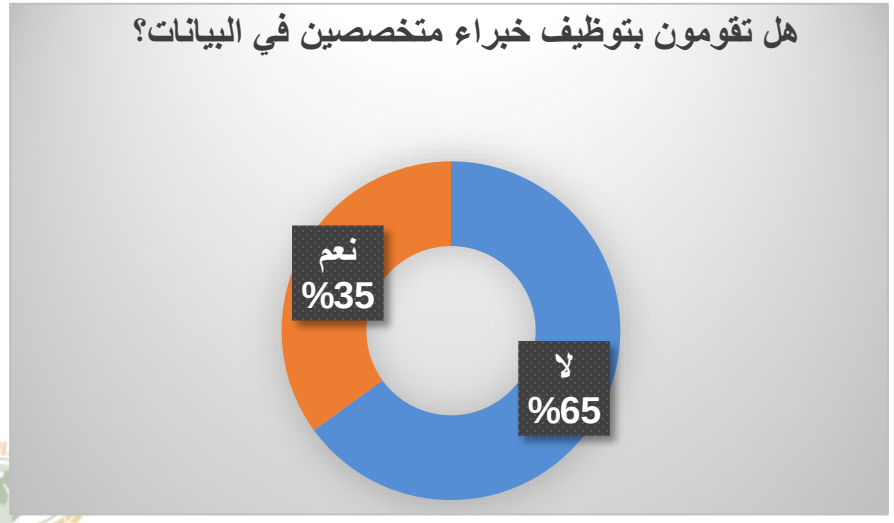
- ترغب كل الجهات المعنية في التعاون مع نظرائها والمساهمة في وضع خارطة طريق /أو خطة للتعاون الإقليمي حول استخدام وتطوير البيانات الكبيرة.

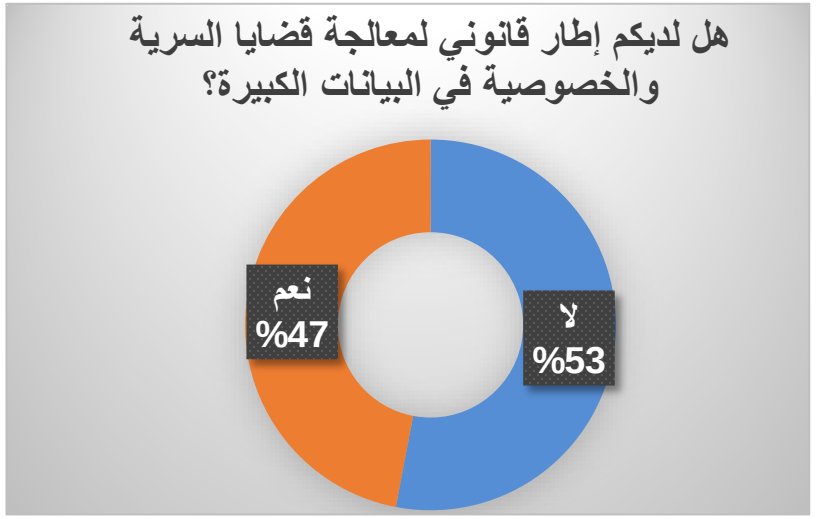
أكبر خمسة تحديات تواجه الجهات المعنية للبدء في مشاريع البيانات الكبيرة

الوضع الراهن – بعض التحديات التي تواجهها الجهات المعنية

- حوالي 41% من الجهات المعنية تستخدم مصادر البيانات الكبيرة.
- حوالي 35% من الجهات المعنية تقوم بتوظيف خبراء مختصين وتدريب موظفيها في مجال البيانات الكبيرة.
- حوالي 53% من الجهات المعنية تفتقد إلى إطار قانوني لمعالجة قضايا سرية وخصوصية البيانات.

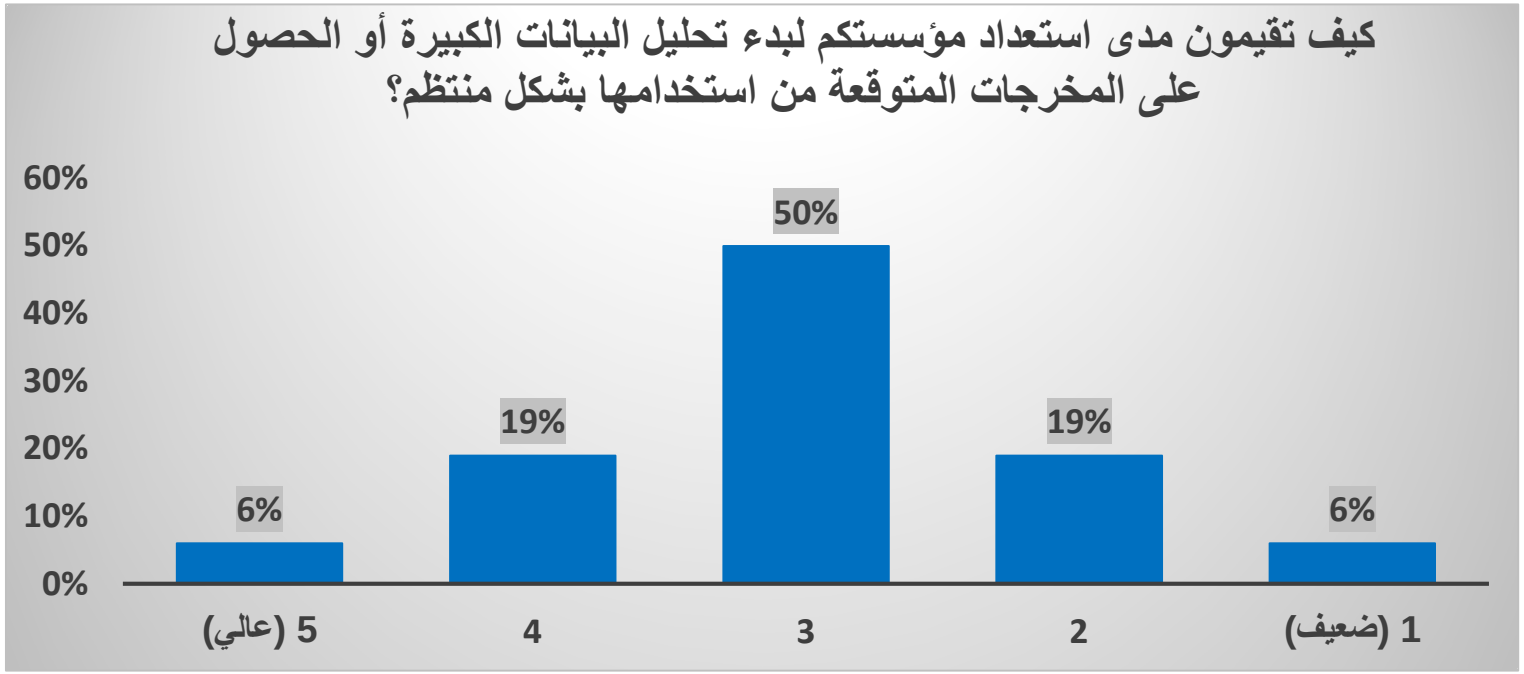






استعداد الجهات المعنية لبدء تحليل البيانات الكبيرة أو الحصول على المخرجات

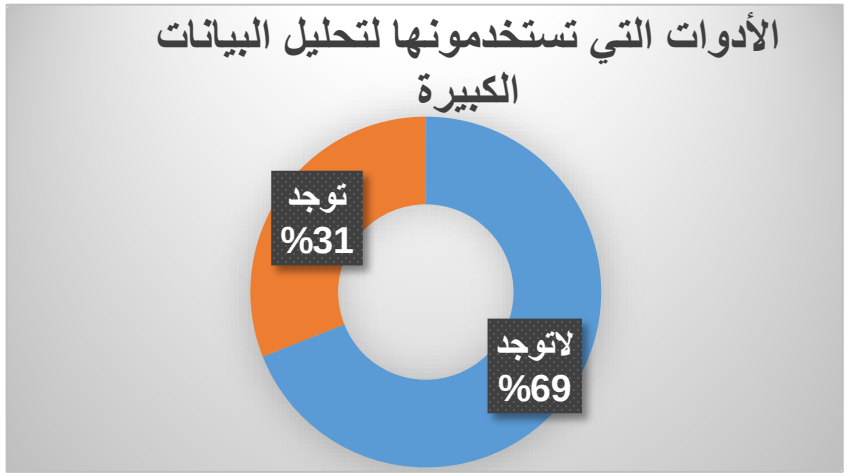
- حوالي 75% من المستطلعين يعتبرون استعداد مؤسستهم بين متواضع وضعيف لبدء عمليات تحليل البيانات الكبيرة أو الحصول على المخرجات المتوقعة من استخدامها بشكل منتظم.

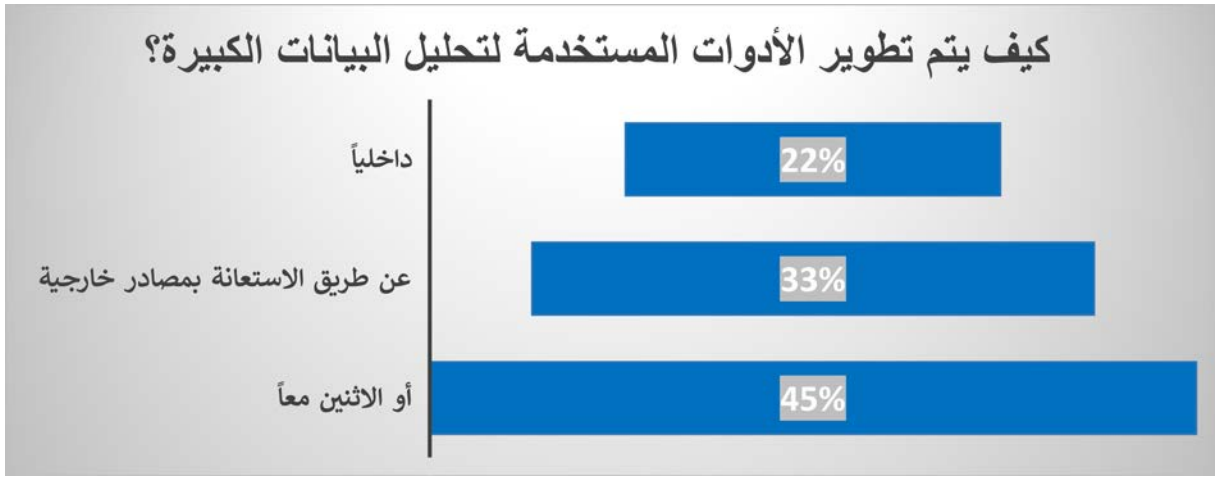


3. غرض ونتائج استبيان تطبيقات البيانات الكبيرة في الدول العربية

استخدام أدوات تحليل البيانات الكبيرة أو الحصول على المخرجات محدود حالياً

- حوالي 69% من المؤسسات العربية المعنية ليس لديها على أدوات لاستخدامها في تحليل البيانات الكبيرة، وحوالي 22% فقط يقوم بتطويرها داخلياً.

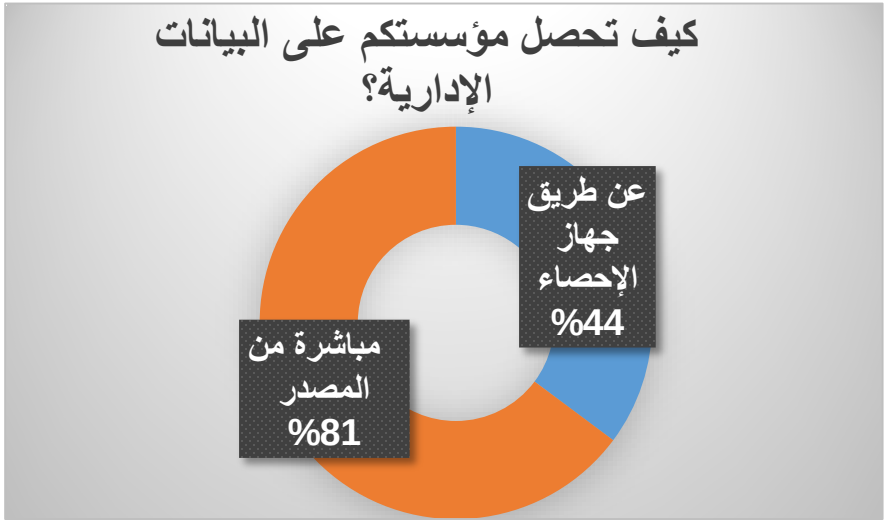


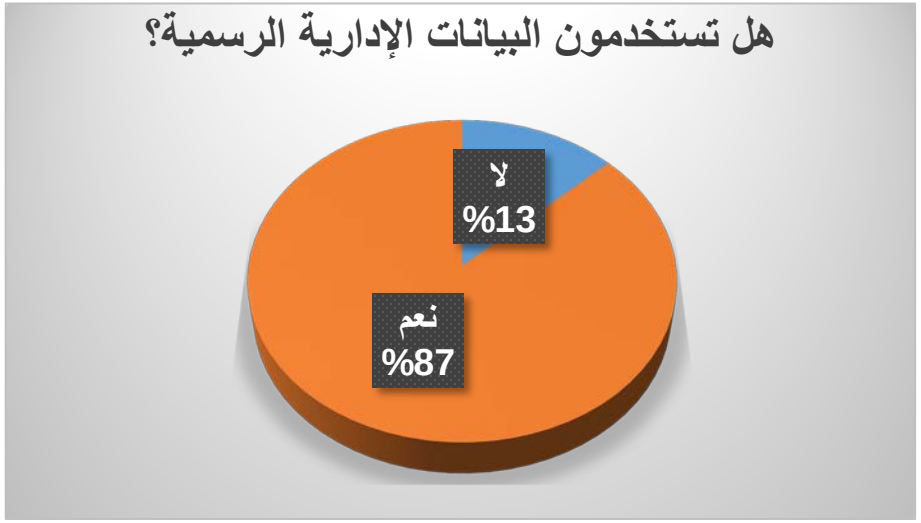


ما الأدوات التي تستخدمونها في تحليل البيانات الكبيرة؟							
"R Language" "R-Studio"	"Alteryx" "QlickView" "KNIME" "Cloudera"	نظام "Fame" المتخصص في السلاسل الزمنية	Microstrategy	Hadoop Postgre SQL ShellScripting	التطلع لاستخدام "Tab lean " كأداة للتحليل وعرض البيانات	بعض التطبيقات الإحصائية المتعلقة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

استخدام البيانات الإدارية

- حوالي 87% من الجهات المعنية تستخدم البيانات الإدارية.
- حوالي 81% من الجهات المعنية تحصل على البيانات الإدارية مباشرة من المصدر.

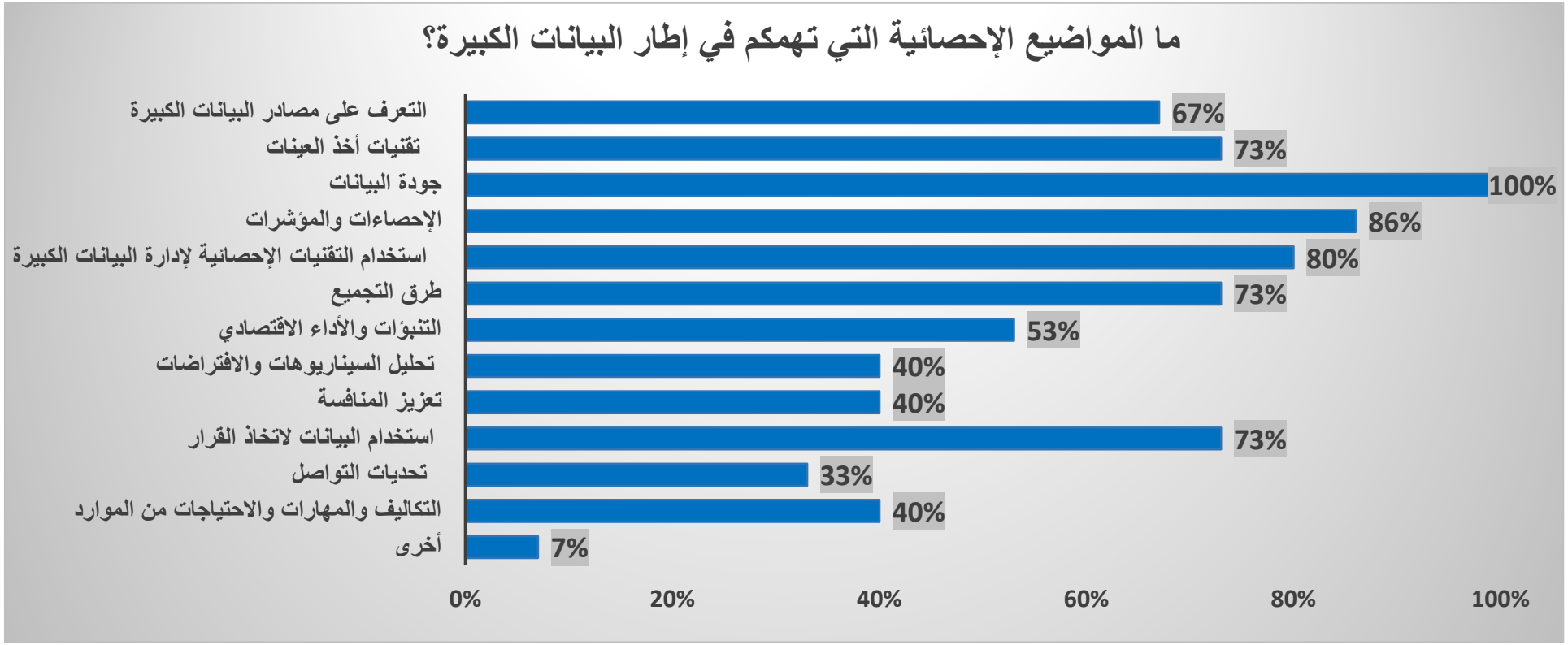




** أكثر من اختيار واحد ممكن لتحديد الجهة المعنية لكيفية حصولها على البيانات الإدارية*

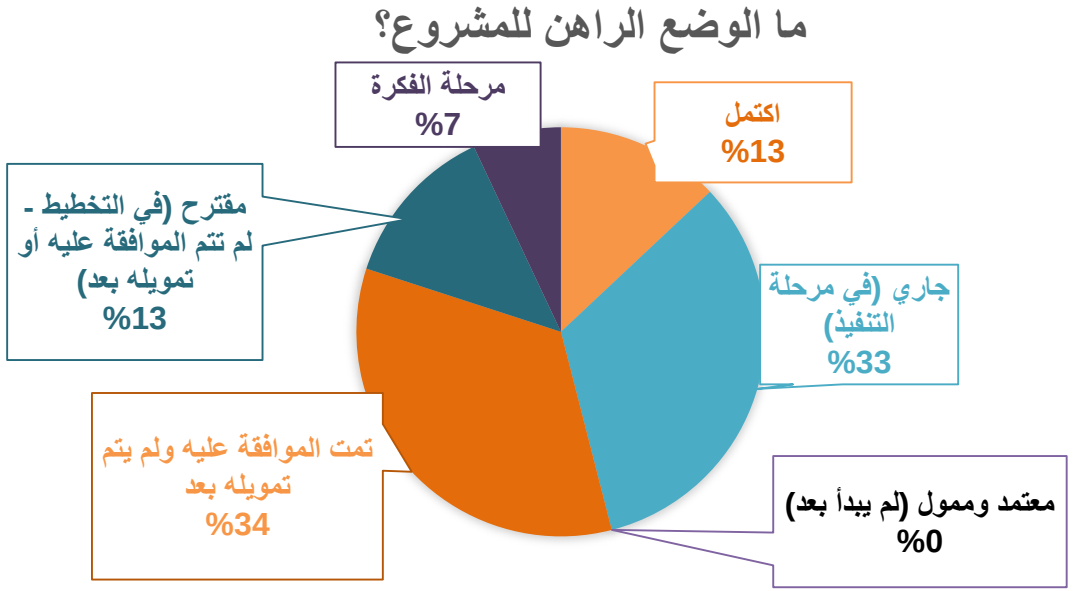
المواضيع الإحصائية التي تهم الجهات المعنية في إطار استخدام البيانات الكبيرة

- تتصدر جودة البيانات تركيز واهتمام الجهات المعنية (100%)، تليها الإحصاءات والمؤشرات (86%).
- حوالي 67% تهتم بالتعرف على مصادر البيانات الكبيرة.

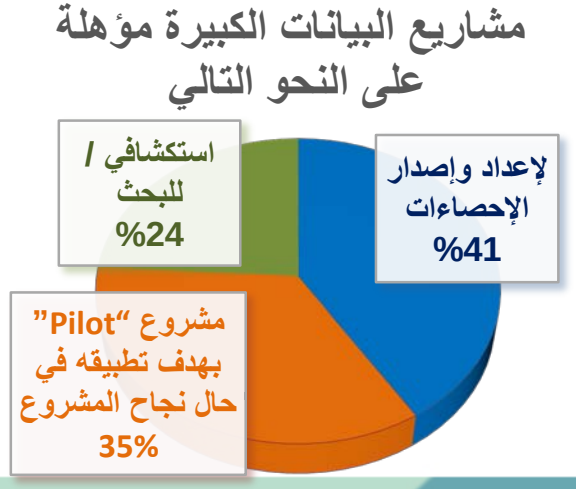


3. غرض ونتائج استبيان تطبيقات البيانات الكبيرة في الدول العربية

مشاريع البيانات الكبيرة محدودة حالياً



- حوالي 34% من مشاريع البيانات الكبيرة تمت الموافقة عليها ولم يتم تمويلها بعد.
- حوالي 13% فقط اكتملت.

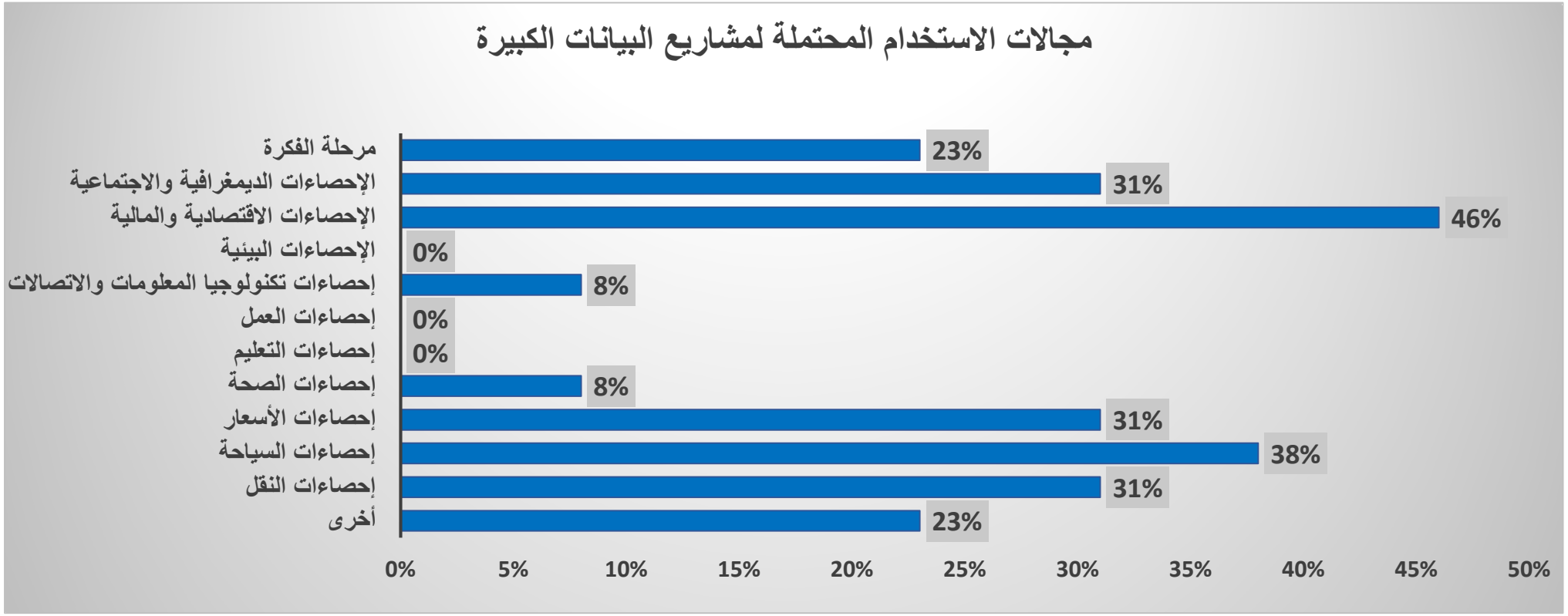


- حوالي 41% من المشاريع مؤهلة لإعداد وإصدار الإحصاءات، و 35% مشاريع "Pilots".

صندوق النقد العربي
ARAB MONETARY FUND

مجالات الاستخدام المحتملة لمشاريع البيانات الكبيرة

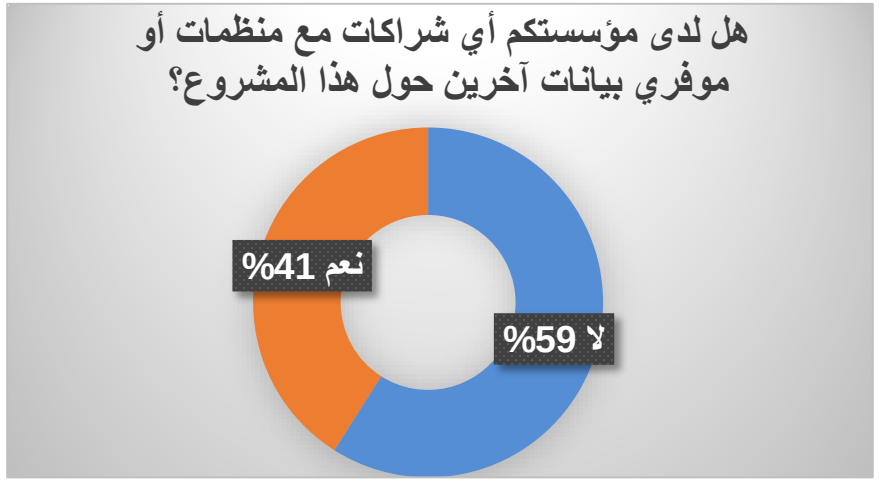
- هناك حوالي 46% من مشاريع البيانات الكبيرة المحتملة لأغراض الإحصاءات الاقتصادية والمالية، و38% لإحصاءات السياحة، و31% لإحصاءات الأسعار وإحصاءات النقل، و31% للإحصاءات الديموغرافية والاجتماعية.

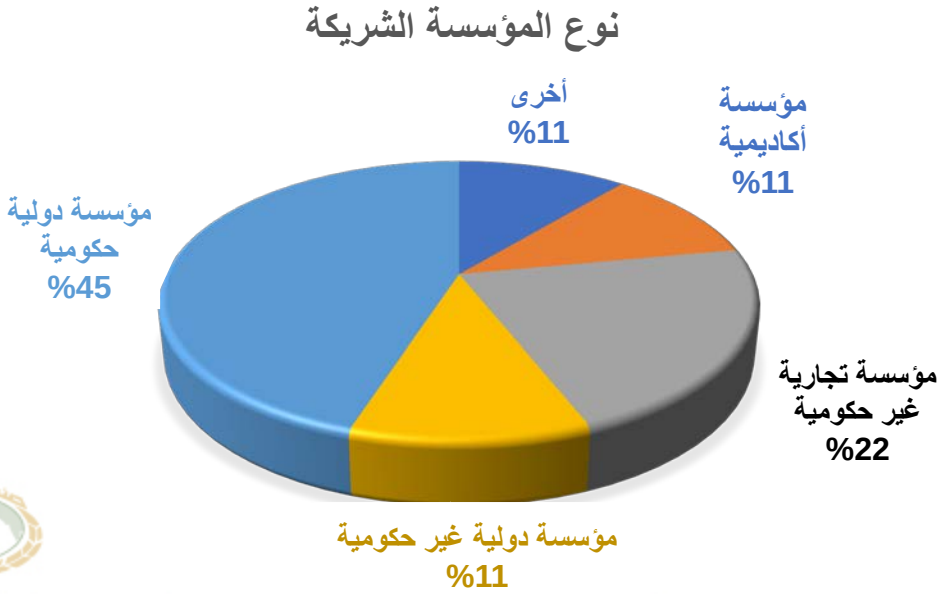


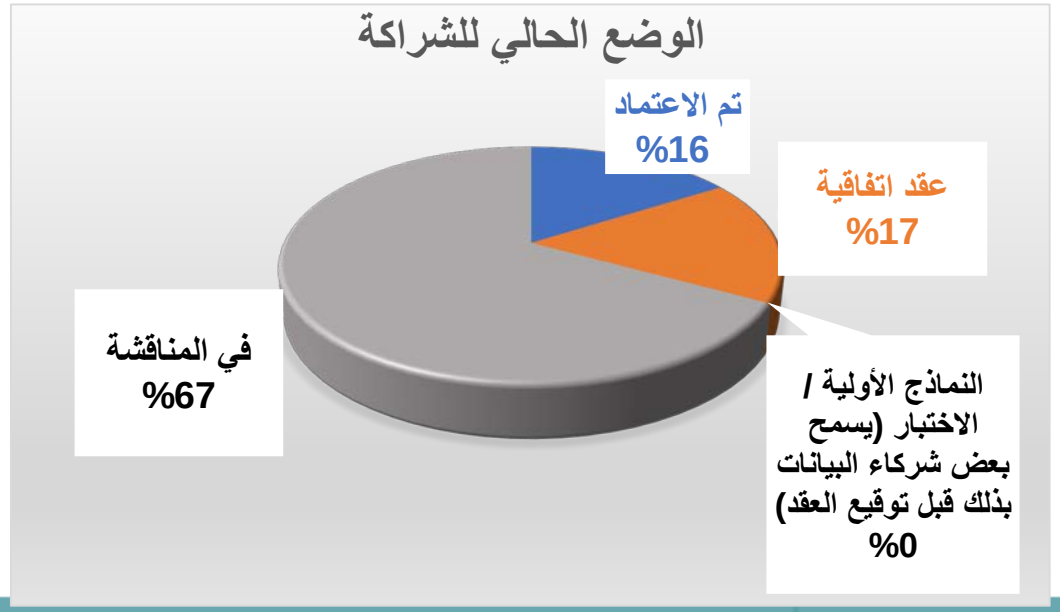
3. غرض ونتائج استبيان تطبيقات البيانات الكبيرة في الدول العربية

الشراكات مع منظمات أو مزودي البيانات حول مشاريع البيانات الكبيرة محدودة

- حوالي 41% فقط من المؤسسات المستطلعة المعنية تتوفر على شراكات مع منظمات أو مزودي بيانات آخرين حول مشاريع البيانات الكبيرة
- حوالي 67% من الشراكات في المناقشة.
- حوالي 45% من الشراكات مع مؤسسات دولية حكومية.







3. غرض ونتائج استبيان تطبيقات البيانات الكبيرة في الدول العربية

مقترحات و آراء الجهات المعنية في الدول الأعضاء

إعداد برنامج متكامل لمساعدة المؤسسات العربية على اكتساب تقنيات البيانات الكبيرة	تنظيم دورات تكوينية منتظمة وبعثات مساندة فنية ومهمات لتبادل الخبرات والكفاءات في مجال البيانات الكبيرة	المساعدة على تطوير البنية التحتية القانونية والتقنية	إنشاء مجموعات وفرق عمل على المستوى العربي والدولي
اتخاذ قرار على مستوى المؤسسات الإحصائية بإدراج موضوع استخدام البيانات الكبيرة ضمن خططها الوطنية نظراً لأهميته الكبيرة	إطلاق المشروعات "Pilots" لتطوير المنهجيات والخبرات المطلوبة وتطوير المنتج الاحصائي الرسمي	نشر الوعي بأهمية الاستفادة من البيانات الكبيرة من خلال مؤتمرات وفعاليات وورش عمل عربية	توفير السياسات المناسبة لتبادل البيانات الكبيرة خصوصاً مع القطاع الخاص
حصر أفضل الممارسات والاستفادة من التجارب الناجحة في مجال استخدام البيانات الكبيرة	المساعدة في توفير موارد مالية مستمرة في إطار تطوير المنظومة المعلوماتية ونظم البيانات الكبيرة	إنشاء بنك الأفكار العربي لمشروعات استخدام مصادر البيانات الكبيرة في الاحصاءات الرسمية وربطها بمشروعات التنمية المستدامة	العمل على بناء علاقات تعاون وشراكات استراتيجية مع مختلف المؤسسات ذات العلاقة

4. بعض تطبيقات البيانات الكبيرة لأغراض الإحصاءات الرسمية وتحليل الاقتصاد الكلي وغيرها

صندوق النقد العربي
ARAB MONETARY FUND

تجربة البنك المركزي الأوروبي في مجال البيانات الكبيرة

دراسات إحصائية مرتبطة باستخدام البيانات الكبيرة	المشاريع " Pilots"
"التنبؤ الآني لنمو الناتج المحلي الإجمالي باستخدام بيانات المدفوعات الإلكترونية"	مجموعة البيانات الإدارية (على سبيل المثال بيانات الميزانية العمومية للشركات)
العلاقات بين التغييرات في ثقة المستهلك ووسائل الاعلام الاجتماعية العامة في هولندا، وإمكانية استخدامها كمؤشر للتغيرات في ثقة المستهلك وكمؤشر مبكر	مجموعة بيانات بحث الويب (مثل معلومات البحث عبر "Google")
تحديد آثار "online bullishness" عبر الانترنت على الأسواق المالية الدولية"، حيث تم تطوير تقدير ثقة المستثمرين (investor sentiment)، استنادا إلى طلبات البحث على "Twitter" و "Google"	مجموعة البيانات التجارية (مثل مشغلي بطاقات الائتمان)
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" .	بيانات السوق المالية (على سبيل المثال تجارة عالية التردد).

المصدر: البنك المركزي الأوروبي

تجربة مكتب الاحصاء الاوروبي (EUROSTAT)

دراسات إحصائية وأنشطة مرتبطة باستخدام البيانات الكبيرة	المشاريع " Pilots"
مطابقة 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.	التجريف على شبكة الإنترنت (Web scraping)، وتشمل الوظائف الشاغرة وخصائص المؤسسات.
تحليل التشريعات والاستراتيجيات، كذلك الأخلاقيات والاتصال (Ethics and communication)، نظراً للتحديات المرتبطة بجودة هذه البيانات.	العدادات الذكية في مجال استهلاك الكهرباء والمساكن المؤقتة.
توفير برنامج تدريبي للإحصائيين الأوروبيين، يشمل: - دورات مخصصة للبيانات الكبيرة - التركيز على مصادر البيانات الكبيرة وعلى أدوات البيانات الكبيرة - اكتساب المهارات اللازمة لتقييم المصادر ونوعيتها، والمهارات اللازمة لاستخدام الأدوات، واستكشاف مصادر البيانات الكبيرة.	مشروع دراسة الجدوى بشأن استخدام بيانات تحديد المواقع المتنقلة في الإحصاءات السياحية (the use of mobile positioning data for tourism statistics).
	بيانات التعرف على أنظمة السفن.

المصدر: مكتب الاحصاء الاوروبي (EUROSTAT)

تجربة المكتب الأمريكي للتحليل الاقتصادي في التعاون مع شركات المعلومات

التعاون مع الشركات
التعاون مع شركة "IBM" لتطوير أسعار الحواسيب الهائلة.
التعاون مع شركة "Chrysler" ومصنعي السيارات الآخرين في تطوير بيانات تسعير السيارات والشاحنات الجديدة.
التعاون مع "Google" لتطوير محركات البحث الإقليمية.

المصدر: المكتب الأمريكي للتحليل الاقتصادي

أنشطة فريق العمل العالمي للأمم المتحدة المعني بالبيانات الكبيرة

2014

- إنشاء فريق العمل العالمي للأمم المتحدة المعني بالبيانات الكبيرة (UN-GWG).

- تقاسم "Pilots" حول مشاريع البيانات الكبيرة.

المؤتمر الأول: بالصين

2015

- الانتقال من "Pilots" إلى المبادئ التوجيهية.

المؤتمر الثاني: بالإمارات العربية المتحدة

2016

- الوصول إلى مصادر البيانات الكبيرة، والشراكات وأهداف التنمية المستدامة

المؤتمر الثالث: بإيرلندا

2017

- التعاون في مجال البيانات والبيانات الموثوقة (Data collaboratives and trusted data)

المؤتمر الرابع: بكولومبيا

تطوير مخزونات البيانات الكبيرة

(Big Data Inventory)

https://unstats.un.org/bigdata/inventor

تطوير بوابة الويب للبيانات الكبيرة

(web gate on big data)

https://unstats.un.org/bigdata

المصدر: لجنة الإحصاء التابعة لمنظمة الأمم المتحدة https://unstats.un.org

البيانات الكبيرة لأغراض الاقتصاد الكلي

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 بألمانيا 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 بالمملكة المتحدة UK unemployment rate & housing market trends	McLaren N & Shanbhogue R (2011)
الاستهلاك الخاص بألمانيا DE private consumption	Vosen & Schmidt (2011)
شراء السيارات بالشيلي Car purchases in Chile	Carriere-Swallow (2011)
التضخم Inflation	Guzmán G (2011)
شراء السيارات بألمانيا German car sales	Fantazzini D & Toktamysova Z (2014)
معدلات البطالة DE, FR, IT, ES NL	Morgan J, e all (2015)

المصدر: الدراسات المذكورة أعلاه

تجربة الحكومة الصينية لتطوير "نظام الائتمان الاجتماعي" باستخدام البيانات الكبيرة

مبادرة تصنيف كل واحد من المواطنين حسب درجة ائتمان باستخدام البيانات الكبيرة: بيانات الهواتف الذكية والتجارة الإلكترونية ووسائل الإعلام الاجتماعية، وبيانات السلطات القضائية (المحاكم) والشرطة والبنوك والسجلات الضريبية وبيانات التوظيف، إلخ.

يمكن لدرجة ائتمان كل مواطن أن تتأثر بشكل سلبي بهوايته وعاداته التسوقية ونمط حياته، وكذلك ما يقوم به من تواصل اجتماعي، مثل، القراءة أو الشراء أو النشر على الإنترنت من تعليقات وآراء سياسية، إلخ.

بعض الأدوات المستخدمة في مجال البيانات الكبيرة

معالجة البيانات Data Cleaning
OpenRefine – DataCleaner

تخزين البيانات وإدارتها
Hadoop - Cloudera - MongoDB - Talend

دمج البيانات Data Integration
Blockspring - Pentaho

تجميع البيانات Data Collection
Import.io

تنقيب البيانات Data Mining
RapidMiner - IBM SPSS Modeler - Oracle data mining - Teradata – FramedData – Kaggle

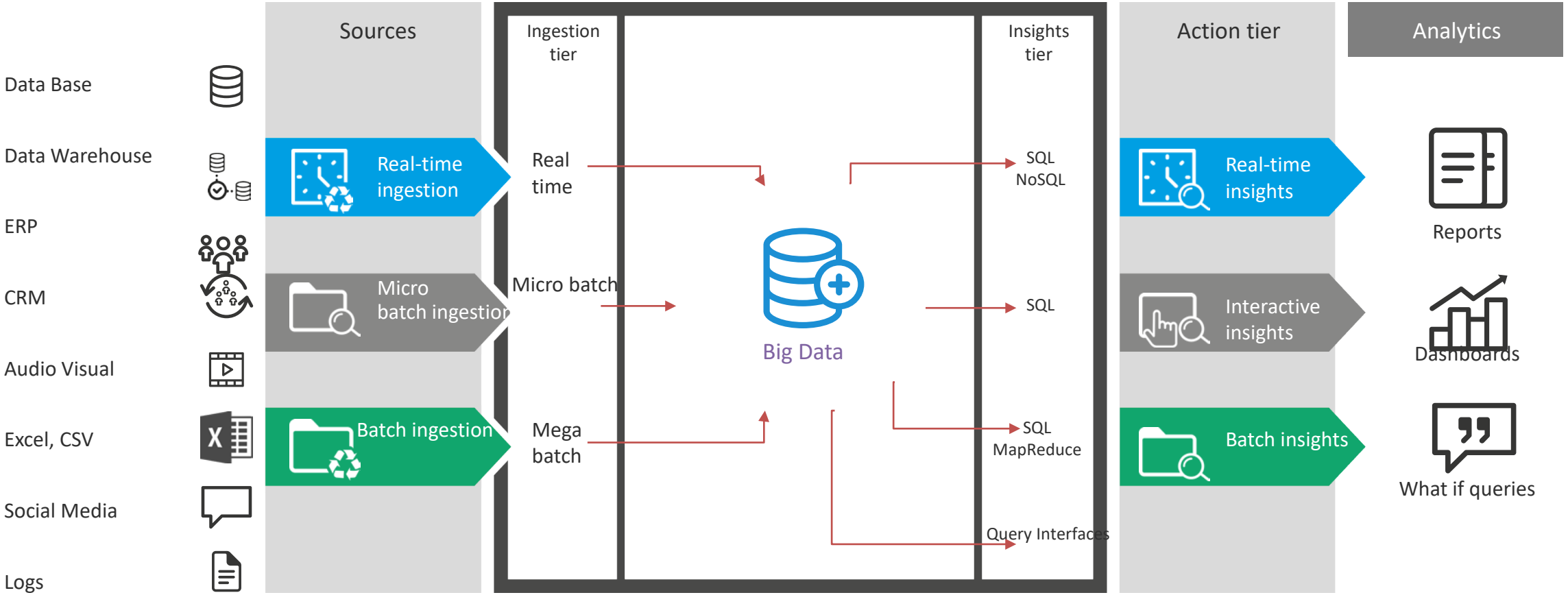
Data Languages
R Language – Python – RegEx – Xpath

العرض المرئي للمعلومات Data Visualization
Tableau – Silk – CartoDB – Chartio - Plot.ly – Datawrapper

تحليل البيانات Data Analysis
Qubole – BigML – Statwing

المصدر: مزودي نظم المعلومات المذكورين أعلاه.

نموذج مشروع البيانات الكبيرة: مركز البيانات أو بحيرة البيانات



مستودع البيانات
"Data warehouse"

vs

مركز أو بحيرة البيانات
"Data Lake or Data Hub"

المصدر: (ITU)International Telecommunication Union

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

صندوق النقد العربي
ARAB MONETARY FUND

المحور الأول - إنشاء فريق عمل إقليمي معني بالبيانات الكبيرة

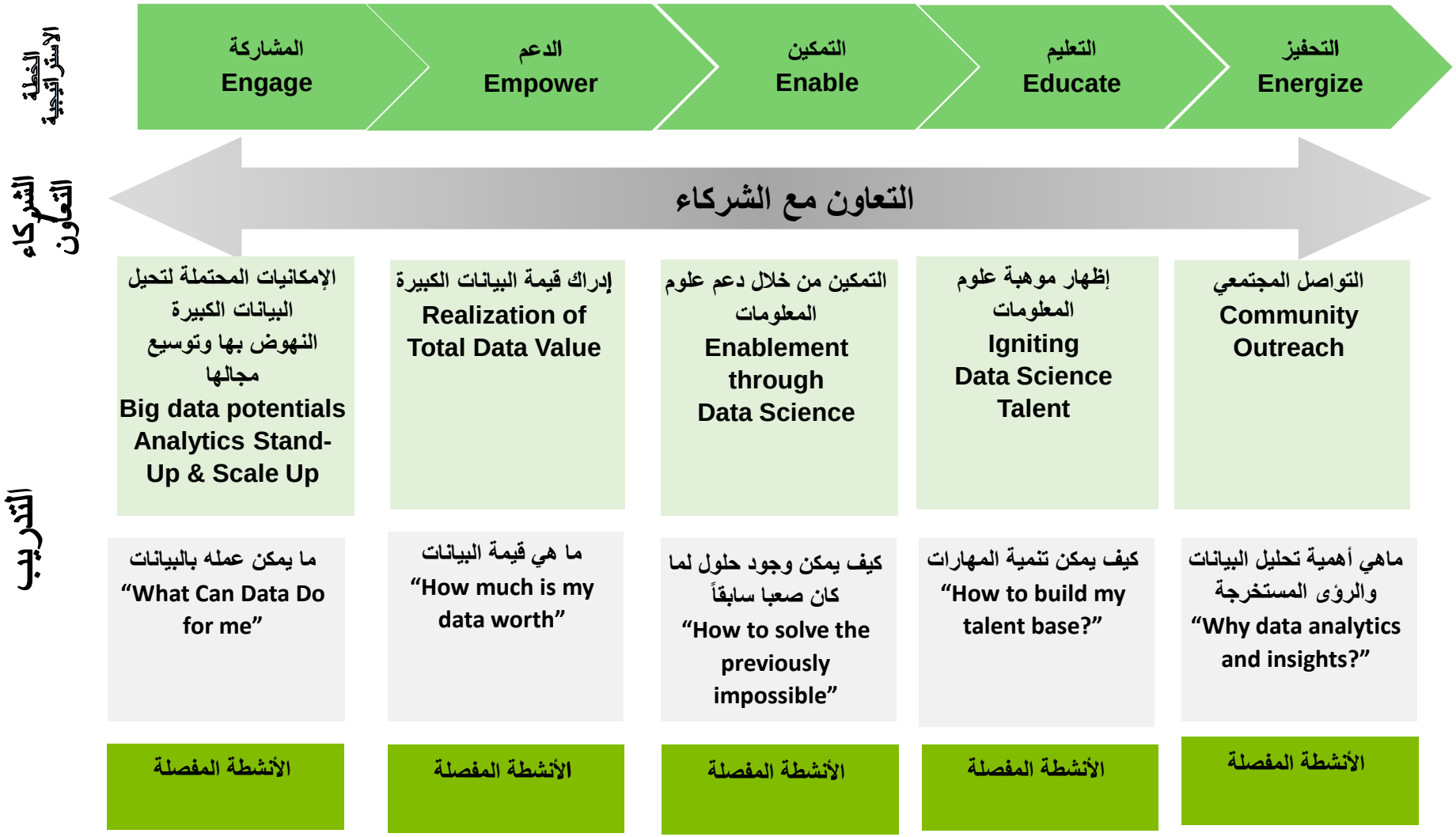
الحوكمة	
الأخلاقيات /الاتصال Ethics / Communication	الجودة Quality
السياسات Policies	التشريعات Legislation
البنية التحتية لتكنولوجيا المعلومات IT Infrastructures	المهارات Skills

مصادر البيانات الكبيرة Big data sources	تبادل الخبرات Experience sharing	المنهجيات Methods

المشاريع "Pilots"

إنشاء فريق عمل إقليمي معني بالبيانات الكبيرة

المحور الثاني - مبادرة إقليمية لبناء القدرات في مجال البيانات الكبيرة



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

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

رصد المعلومات وأدوات جديدة، ومراجع، ودورات تدريبية، وتبادل الخبراء....

تقييم جاهزية واستعداد البلدان الأعضاء من خلال استبيان تطبيقات البيانات الكبيرة في البلدان العربية (صندوق النقد العربي)

" مركز إقليمي لتبادل المعرفة في مجال البيانات الكبيرة"

Regional

Big Data Excellence Centre

تقديم الأنشطة المناسبة للبلدان الأعضاء

التنسيق والاتصال

استخدام الموارد المتاحة

الشراكات مع المصادر الأخرى ومقدمي الخدمات

مطابقة المهارات بين الدول الأعضاء والمبادرات الإقليمية والعالمية

شكراً

صندوق النقد العربي

http://www.amf.org.ae