ملتقى المعلومات الصناعية والإحصاء في الدول العربية

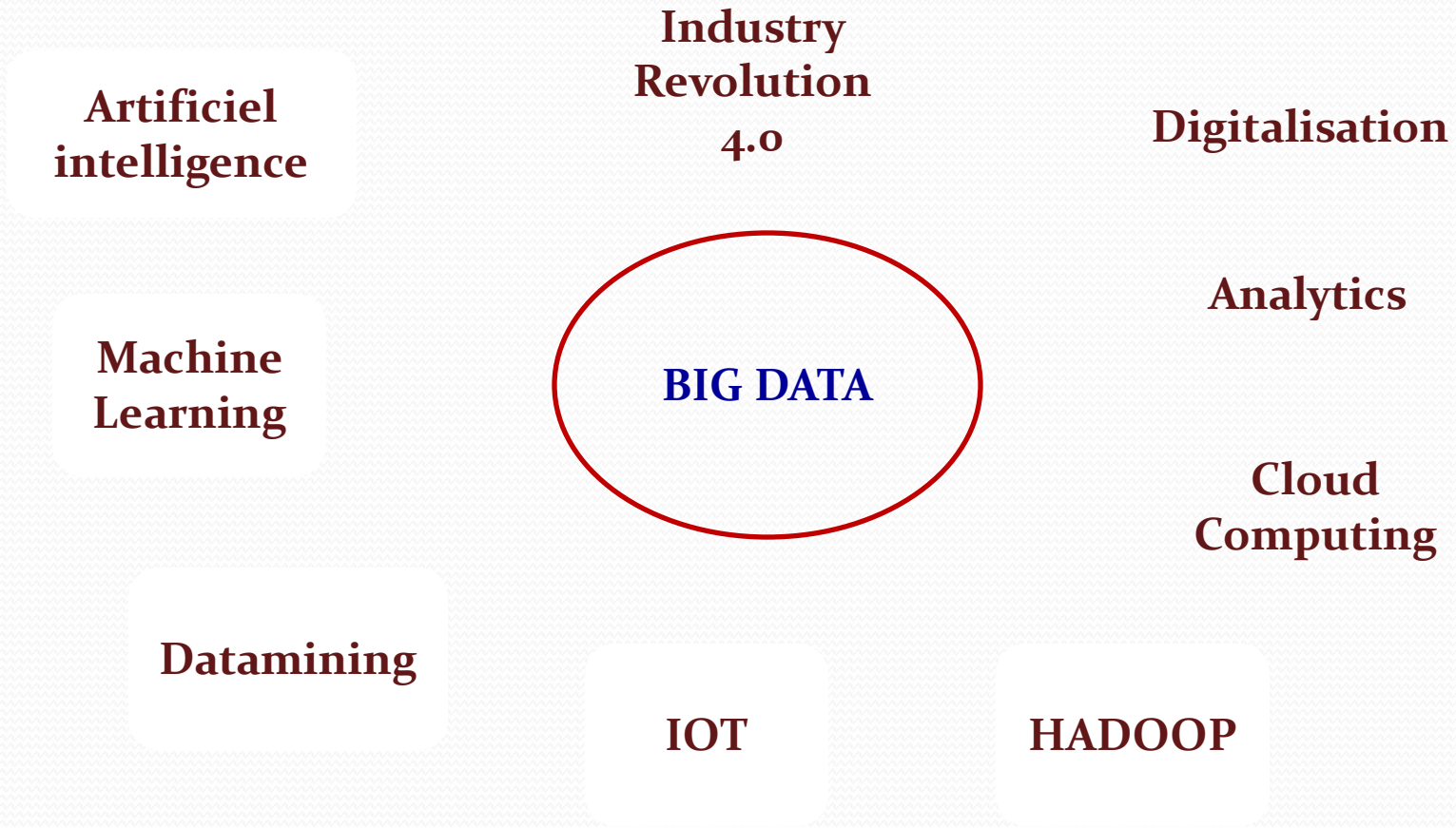
الرباط 7-8 نوفمبر 2018

البيانات الضخمة والذكاء الصناعي

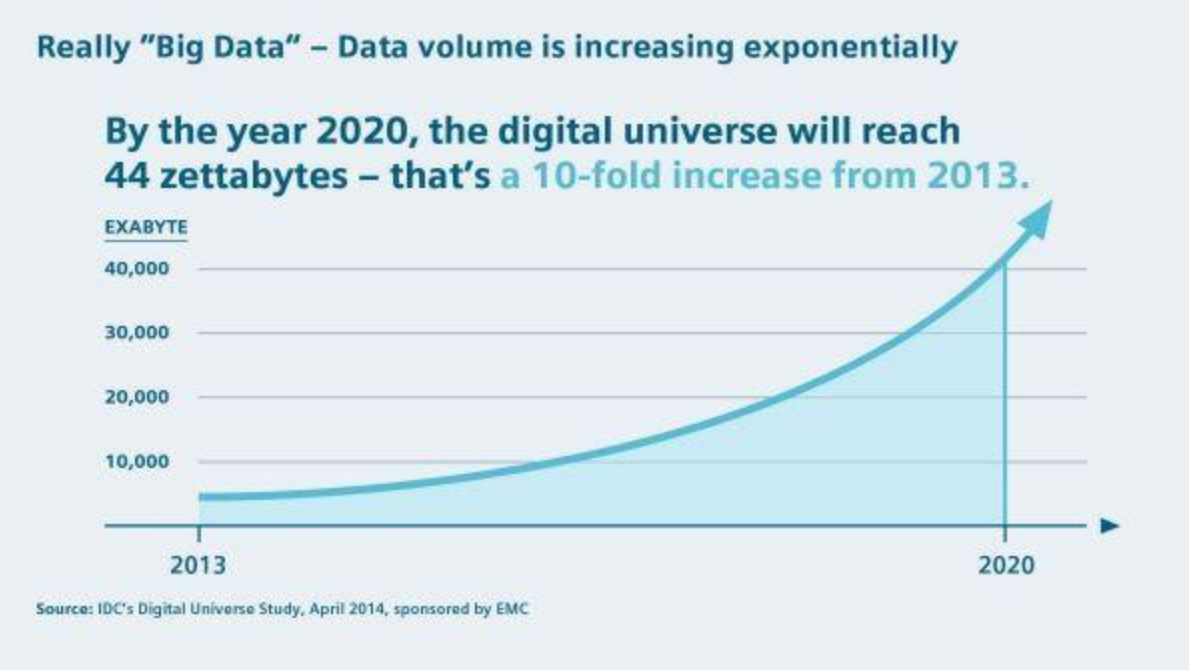
الدكتورة بسمة بوكنز

كلية العلوم والتقنيات ، جامعة الحسن الأول، سطات

المقدمة (ثورة البيانات)



Technological Development



جدول المحتويات

البيانات الضخمة

1)مفهوم البيانات الضخمة

2) خصائص البيانات الضخمة

3)مصادر البيانات الضخمة

أهميتها وأهمية الاستثمار في تحليلها

3) تقنيات البيانات الضخمة(BIG DATA Technics)

التنقيب عن البيانات (Datamining)

الذكاء الإصطناعي

البيانات الضخمة

البيانات الضخمة (BIG Data)

هي مجموعة ضخمة من البيانات المعقدة

و المتداخلة بشدة

(كالتغريدات على تويتر، الدردشة، الرسائل النصية، الاعجابات بمنتج او منشور معين، مشاركة الحالة أو فيديو مع الأصدقاء، حجم تداول الأسهم، مؤشرات وأخبار الطقس، وغيرها الكثير جدا)

البيانات الضخمة

مصادرالبيانات الضخمة

الخدمات الإلكترونية

E-services.

الهواتف النقالة

Mobiles

وسائل التواصل الاجتماعي

Social media

التسوق أونلاين

Online shopping

التلفزيونات الذكية

Smart TV

مواقع الويب

Website

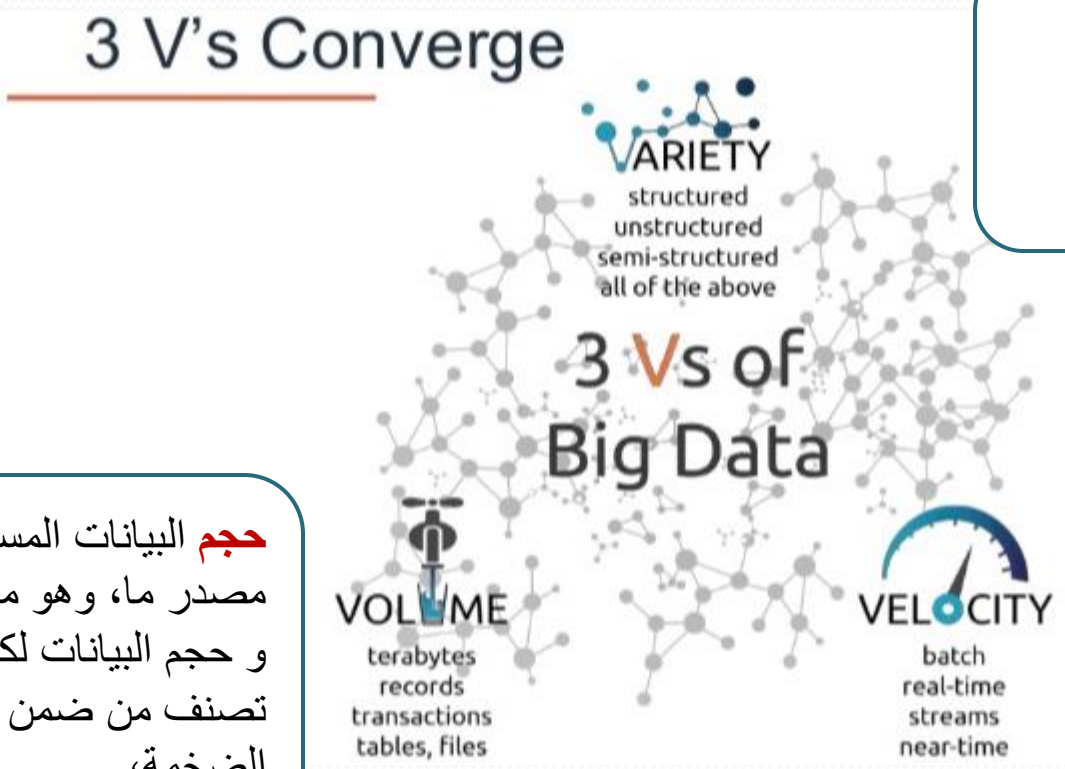
أجهزة الاستشعار

Sensor technology

البيانات الضخمة

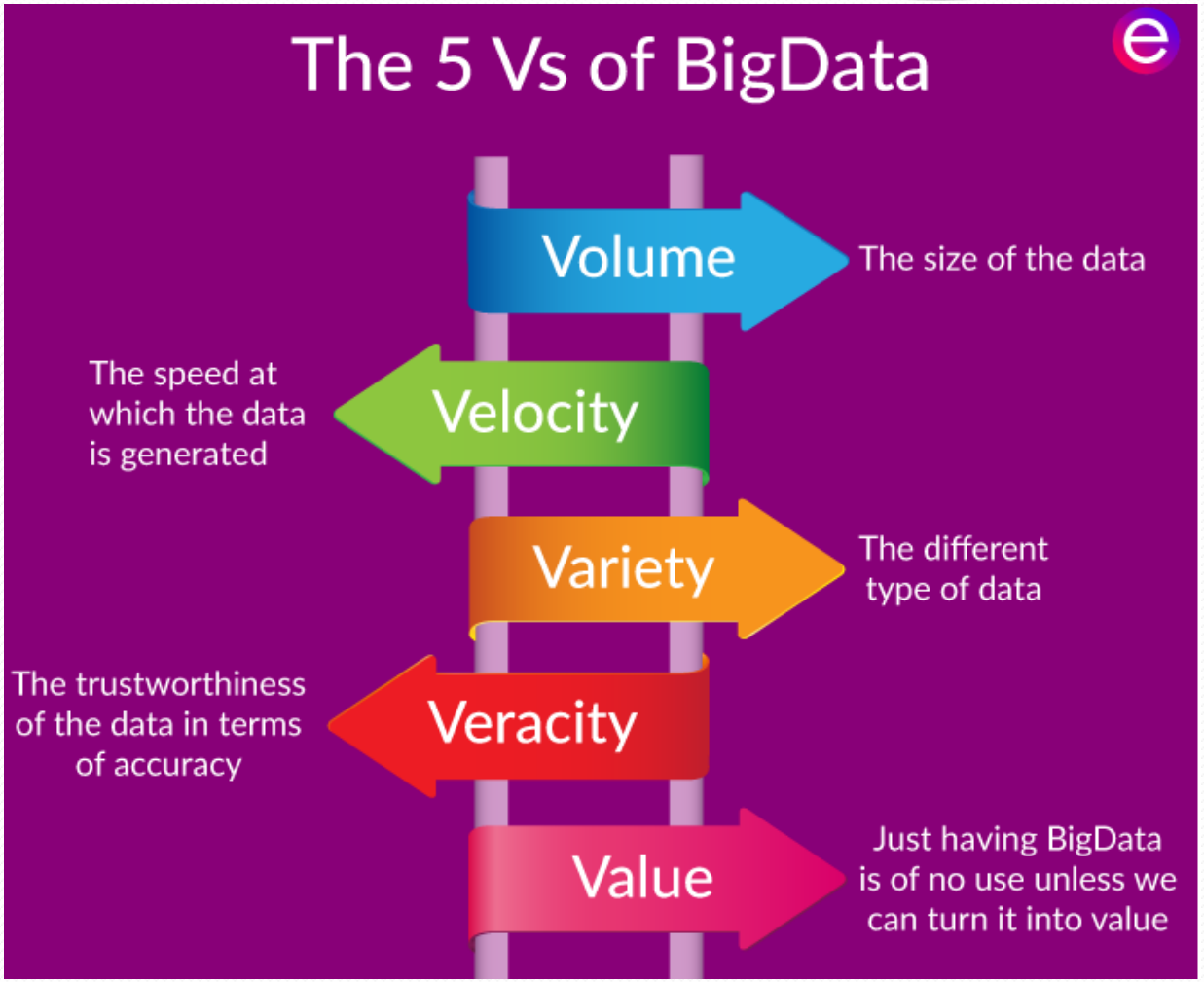
خصائص البيانات الضخمة

يُقصد بها **تنوع** البيانات المستخرجة، والتي تساعد المستخدمين سواء كانوا باحثين أومُحللين على إختيار البيانات المناسبة لمجال بحثهم



حجم البيانات المستخرجة من مصدر ما، وهو ما يحدد قيمة و حجم البيانات لكي تصنف من ضمن البيانات الضخمة،

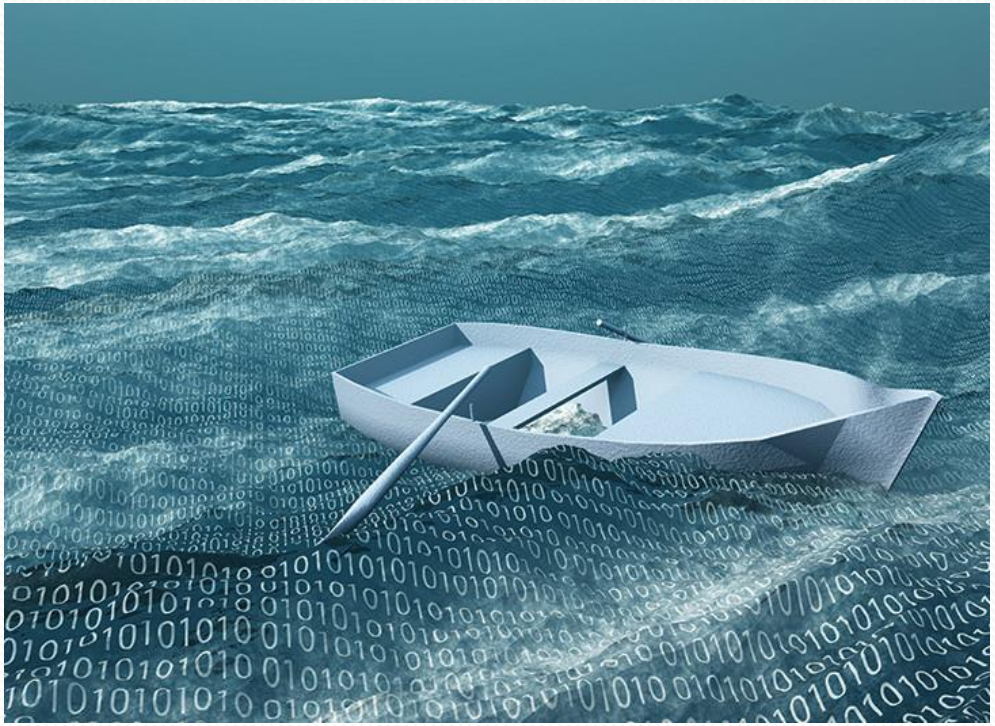
يُقصد بها **سرعة** إنتاج وإستخراج البيانات لتغطية الطلب عليها حيث تعتبر السرعةعنصراحاسما في اتخاذ القرار بناء على هذه البيانات



موثوقية مصدر البيانات، ومدى دقتها وصحتها وحداثة تلك البيانات الضخمة،

القيمة

أهميتها؟



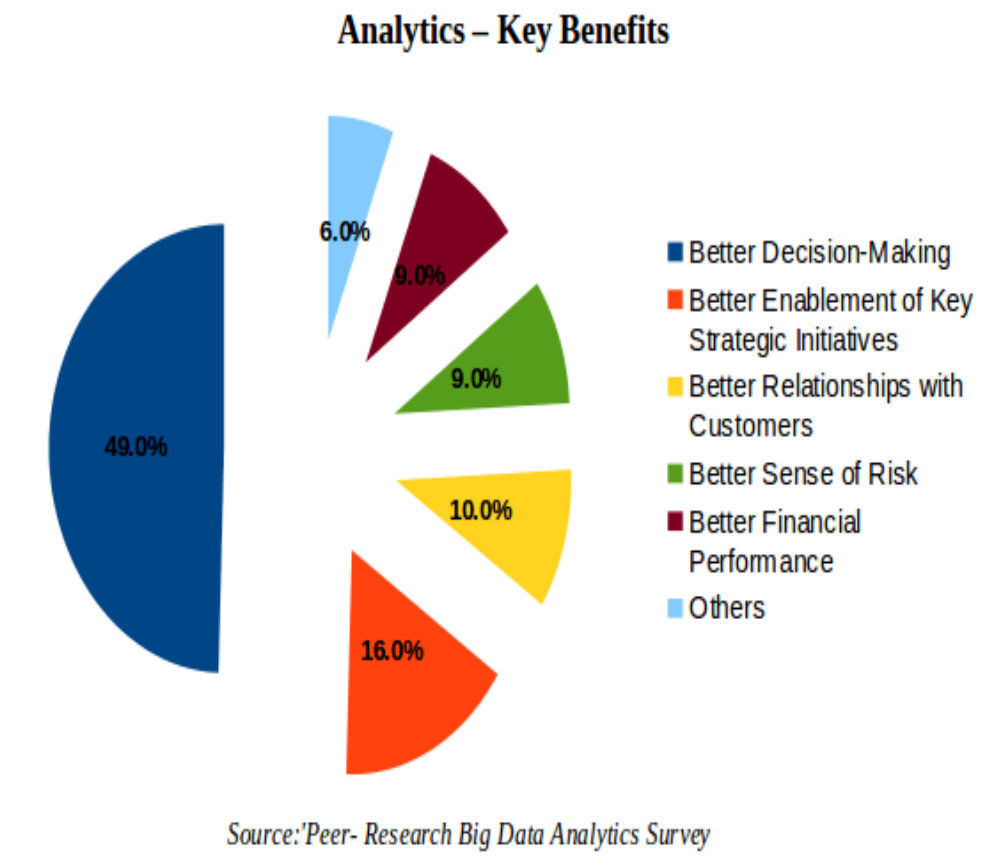
البيانات الضخمة

BIG DATA

Analytics/Predictive Analytics/Knowledge Discovery

البيانات الضخمة

BIG DATA



وأهمية الاستثمار في تحليلها؟

تحسين العمليات الداخلية، مثل إدارة المخاطر، إدارة علاقات العملاء، والخدمات اللوجستية

تحسين المنتجات والخدمات القائمة.

- تطوير الخدمات والمنتجات الجديدة.

- ا لإستفادة من المعلومات وتقديم العروض المناسبة للعملاء في الوقت المناسب

تطبيقات البيانات الضخمة وأثرها

يمكن للحكومات تحليل محتوى مواقع التواصل الاجتماعي لمواطنيها حول قراراونظام معين مطبق او تريد تشريعه وتطبيقه وبالتالي معرفة ردود الافعال حول ذلك.

يمكن للشركات الإستفادة من تحليل البيانات الناتجة من وسائل التواصل

الإجتماعي مثل فيسبوك وتويتر لتحديد جمهورها والتنبؤ بنتائج حملة التسويق والمبيعات

الأحوال الجوية الإستفادة من المعلومات المتوافرة فيها لاستخلاص تنبؤات جوية فائقة الدقة

تطبيقات البيانات الضخمة وأثرها

التنبؤ بالكوارث الطبيعية

تستهدف فيسبوك الفئات التي توجه لها الإعلانات بناء على تحليل إهتمامات مستخدميه وطبيعتهم،

تحليل البيانات الضخمة وتحسين التعليم

تحليل البيانات الضخمة لتحسين عملية صنع القرار

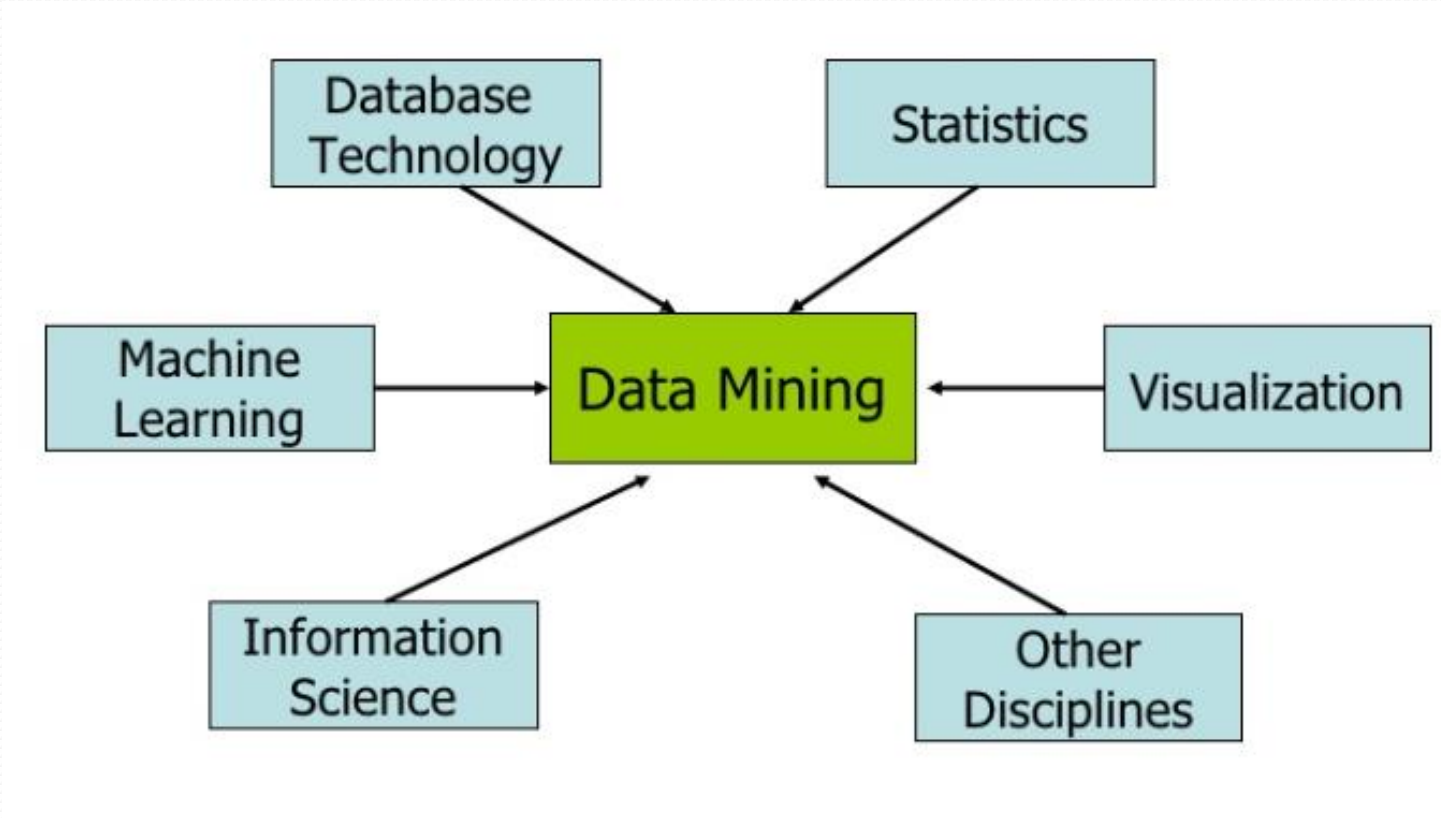
الإستفادة من البيانات الضخمة في المجال العسكرى

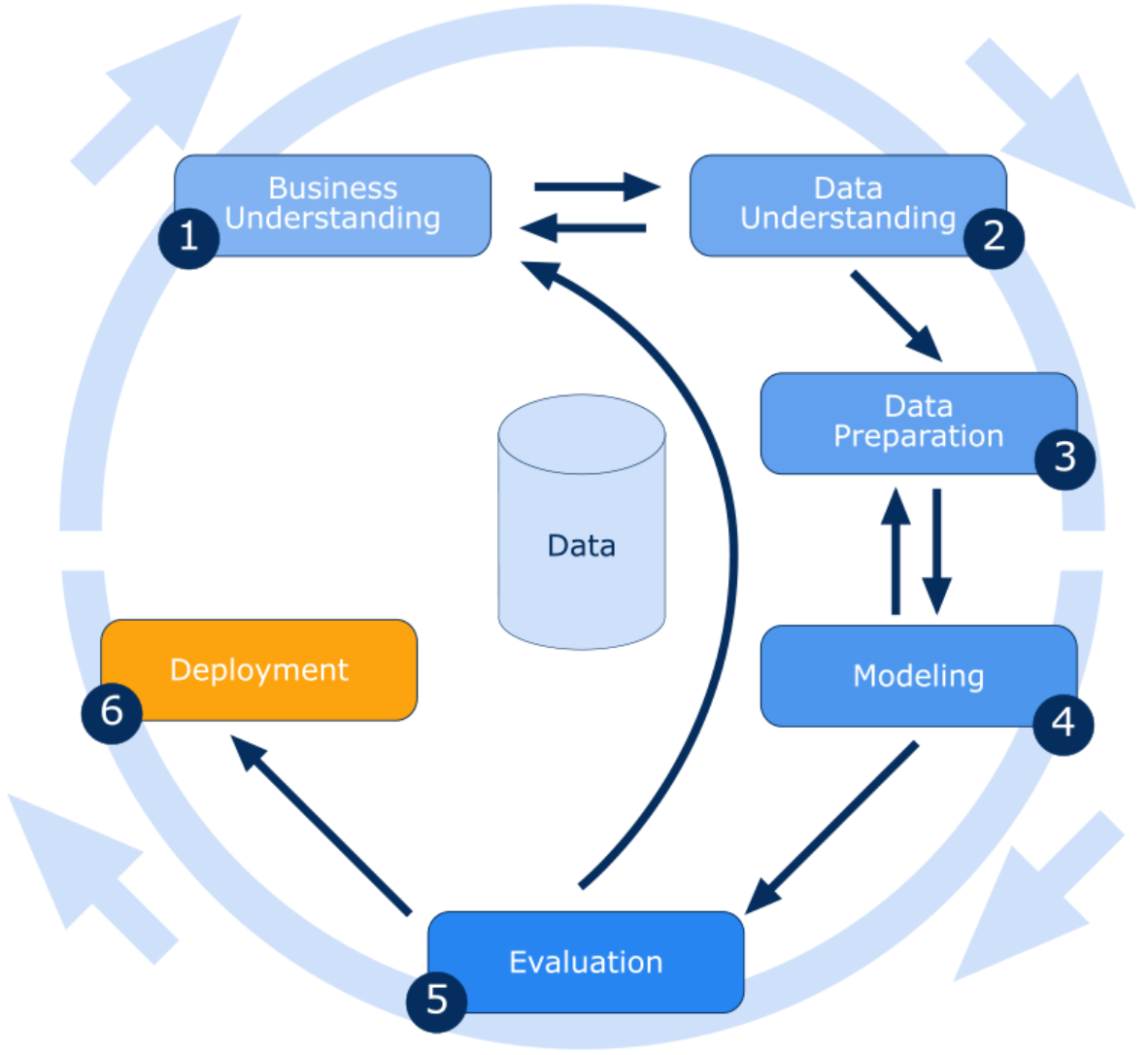
الإستفادة من البيانات الضخمة في المجال الإقتصادى

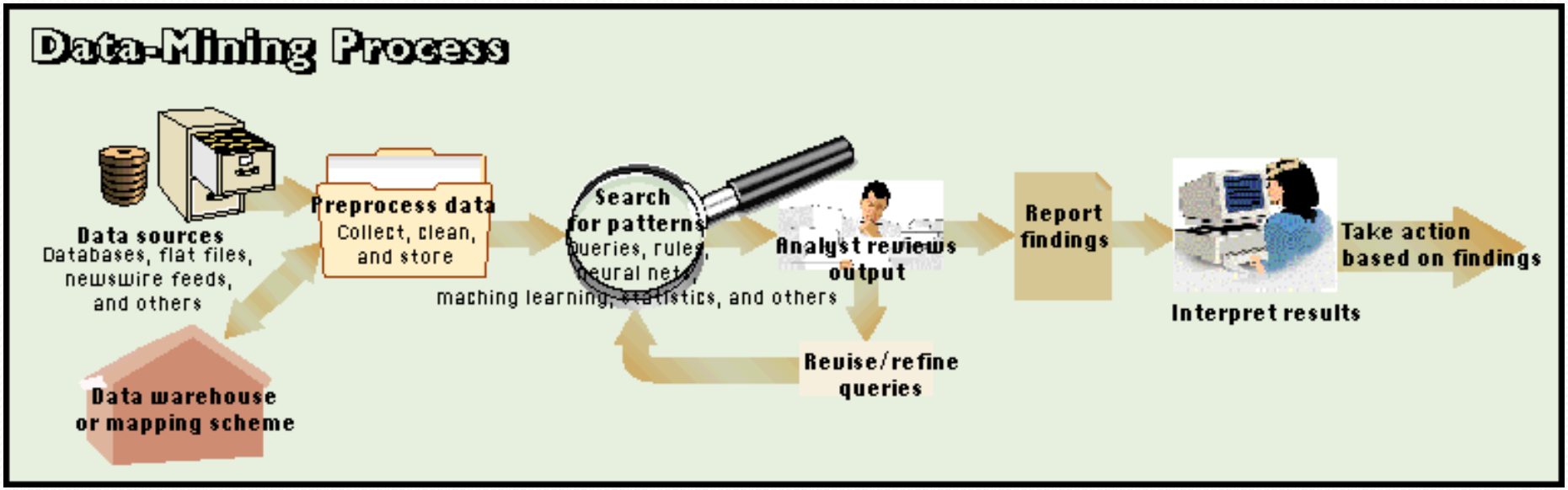
الإستفادة من البيانات الضخمة في المجال الطبى

التنقيب عن البيانات
(Datamining)

كتقنية تهدف إلى استنتاج المعرفة من كميات هائلة من البيانات، تعتمد على الخوارزميات الرياضية والتي تعتبر أساس التنقيب عن البيانات وهي مستمدة من العديد من العلوم مثل علم الإحصاء والرياضيات والمنطق وعلم التعلم، والذكاء الاصطناعي والنظم الخبيرة، وعلم التعرف على الأنماط ،وعلم الآلة. وغيرها من العلوم والتي تعتبر من العلوم الذكية وغير التقليدية.


Database Technology
Statistics
Machine Learning
Data Mining
Visualization
Information Science
Other Disciplines


Business Understanding
1
Data Understanding
2
Data Preparation
3
Modeling
4
Evaluation
5
Deployment
6
Data


Data-Mining Process
Data sources
Databases, flat files,
newswire feeds,
and others
Preprocess data
Collect, clean,
and store
Search
for patterns
Queries, rules,
neural nets,
machine learning, statistics, and others
Analyst reviews
output
Report
findings
Interpret results
Take action
based on findings
Revise/refine
queries
Data warehouse
or mapping scheme

الذكاء الاصطناعي

الذكاء الاصطناعي

ماهية الذكاء الإصطناعي

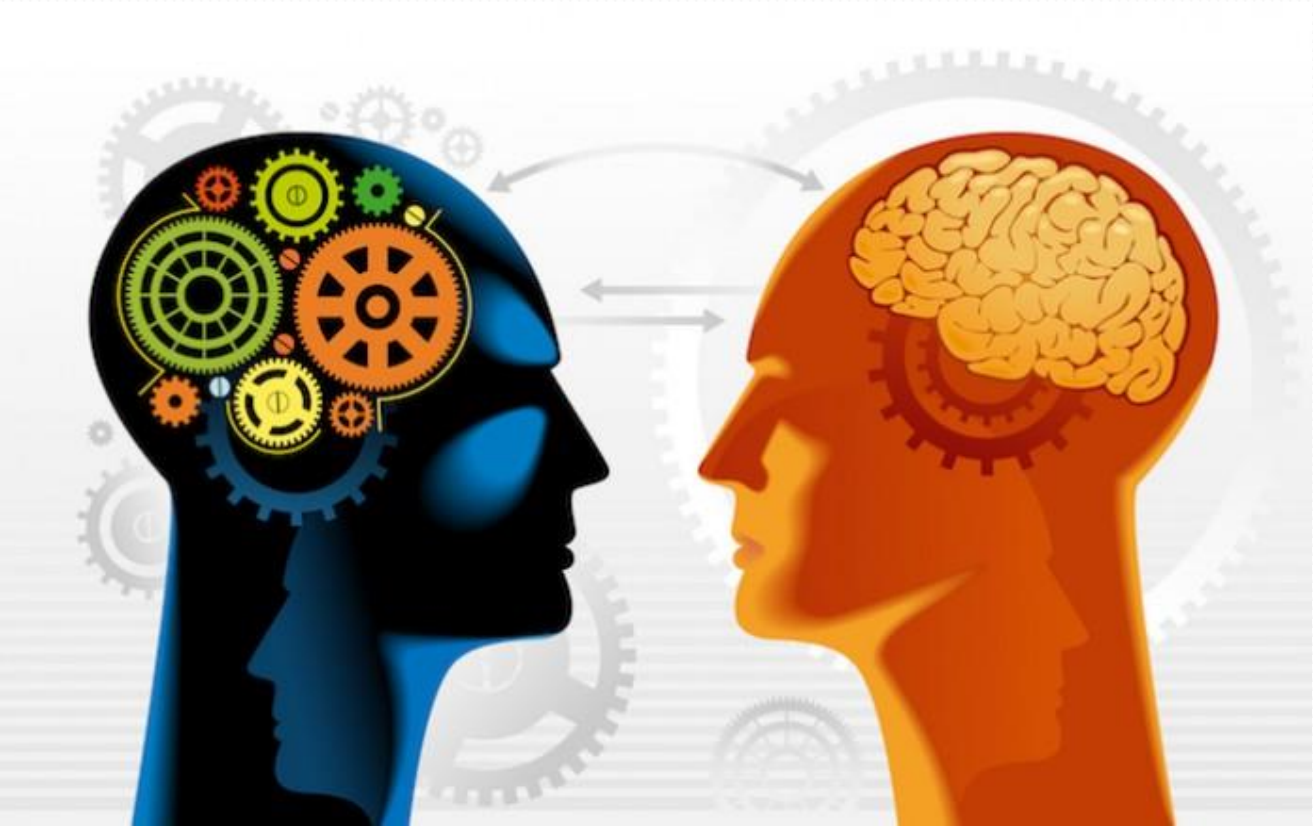
لمعرفة ماهية الذكاء الاصطناعي يتعين أولا تحديد المقصود بالذكاءالانساني

يرتبط بالقدرات العقلية مثل القدرة على التكيف مع ظروف الحياة والاستفادة من التجارب والخبرات السابقة والتفكير والتحليل والتخطيط وحل المشاكل والاستنتاج السليم والاحساس بالآخرين ،

بالإضافة إلى سرعة التعلم واستخدام ما تم تعلمه بالشكل السليم والمفيد.

الذكاء الاصطناعي

أما الذكاء الاصطناعي فهو محاكاة لذكاء الانسان وفهم طبيعته عن طريق عمل برامج للحاسب الآلي قادرة على محاكاة السلوك الانساني المتسم بالذكاء



الذكاء الاصطناعي

السيارات ذاتية القيادة والطائرات بدون طيار.

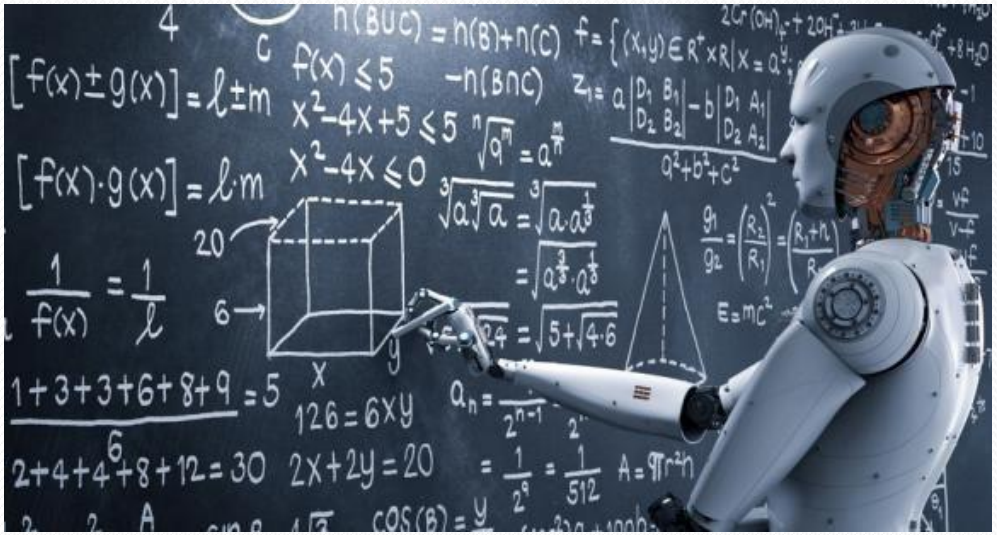
الانسان الآلي (الروبوت)

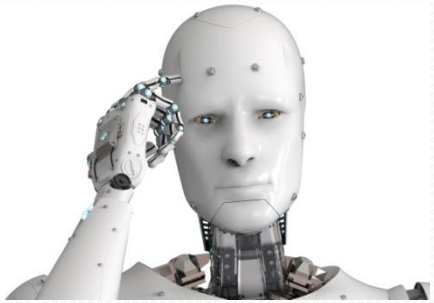
خدمات المنازل الذكية ، والأسلحة ذاتية العمل ، والهواتف الذكية ، وأجهزة التلفاز الذكية ،

الأجهزة الذكية القادرة على القيام بالعمليات الذهنية كفحص التصاميم الصناعية ، ومراقبة العمليات واتخاذ القرار.

الذكاء الاصطناعي

- في العام 1958 اخترع العالم " جون مكارثي " لغة البرمجة(" lisp ") اللغة المفضلة في مجال الذكاء الاصطناعي





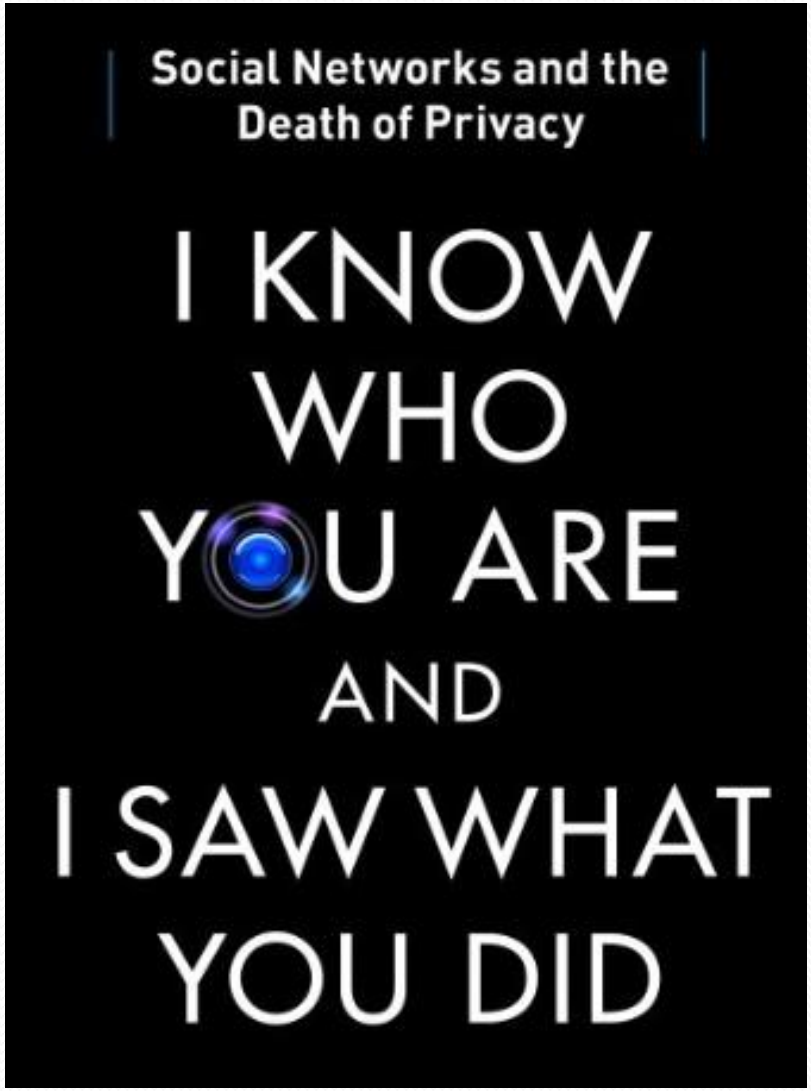
vidéo

التحديات

هناك بعض التحديات الجديدة المرتبطة بالبيانات الضخمة، منها:

- الحصول على بيانات الحكومة والقطاع الخاص، فضلاً عن الموارد الحاسوبية اللازمة.
- تمثّل معالجة أحجام من البيانات سريعة الازدياد تحدياً صعباً أيضاً، مع تزايد معدلات تدفق البيانات بسرعة هائلة.
- عدم التجانس وعدم الاكتمال: يمكن أن تفوّت البيانات التي تمّت معالجتها بتقانات البيانات الضخمة بعض سمات البيانات أو تُدخل ضوضاء في أثناء نقلها حتى بعد تنظيف البيانات وتصحيح الخطأ، ومن المرجح أن يظلّ هناك بعض النقص والأخطاء فيها، ويمكن أن تُدار هذه التحديات خلال تحليل البيانات.
- الحجم: تنامي حجم البيانات أسرع من تنامي التقانات،
- الخسارة المالية والسمعة نتيجة خروقات البيانات الضخمة.

التحديات


Social Networks and the
Death of Privacy
I KNOW
WHO
YOU ARE
AND
I SAW WHAT
YOU DID

The GDPR

The General Data Protection Regulation

وهي واحدة من أهم التشريعات واسعة النطاق التي تتعلق بالتكنولوجيا والإنترنت.
هذه اللوائح تم اعتمادها من قبل الاتحاد الأوروبي في أبريل من عام 2016, وتم اتخاذ قرار بدخولها حيز التنفيذ في يوم 25 مايو 2018,
وهي قوانين يجب أن تلتزم بها كافة الشركات والنشاطات التجارية والتطبيقات والمنصات على الإنترنت داخل أوروبا أو التي تستهدف الجمهور من دول الاتحاد الأوروبي.

.

The GDPR

The General Data Protection Regulation

تتعلق هذه القوانين واللوائح بحقوق الخصوصية للمستخدمين وبياناتهم التي ينشئونها عبر الإنترنت.,
ستحتاج الشركات إلى تقديم إشعارات صريحة للمستخدمين , تخبرهم فيها بأنها تقوم بجمع البيانات الشخصية لهم

كما يحق للمستخدم أن يعرف بالتفصيل البيانات التي تحتفظ بها الشركة أو المؤسسة التي يتعامل معها عبر الإنترنت, كما يحق له طلب حذف أي من هذه المعلومات إذا شعر بإنتهاك لحقوقه في الخصوصية.
وأي موقع أو منصة تتعرض بياناتها للاختراق, عليها فوراً ابلاغ السلطات المعنية في غضون 72 ساعة من حدوث ذلك.

WHAT WE KNOW...
THE REST...
BIG DATA

شكراً

البيانات الضخمة

تقنيات البيانات الضخمة(HADOOP)

وهو برنامج أو منصة برمجية مفتوحة المصدر مكتوبة بلغة الجافا لتخزين و معالجة البيانات الضخمة بشكل موزع مثل تخزين بيانات ضخمة على عدة أجهزة ومن ثم توزيع عملية المعالجة على هذه الأجهزة لتسريع نتيجةالمعالجة.

Amazon, Apple, eBay, Electronic Arts, Facebook, Google, IBM, LinkedIn, Microsoft, the New York Times, Twitter, Yahoo

البيانات الضخمة

تتكون الأدوات التي تتعامل مع البيانات الضخمةمن ثلاثة أجزاء رئيسية وهي:

أدوات التنقيب عن البيانات (Datamining)

أدوات التحليل (Data Analysis)

أدوات عرض/تصوير النتائج (Dashboard /Visualization)

ثورة البيانات

مقدمة :

أدى الانتشار الواسع لتقنية المعلومات وسهولةإتاحتها إلى تضخم حجم المعلومات بصورة استباقية لم يشهدها التاريخ من قبل، فلم يسبق للبشرية ان راكمت مثل هذا الكم من البيانات بأي شكل من الأشكال.

ثورة البيانات

كما تمتلك تقنية البيانات الضخمة إمكانية التحليل ,حيث أن تحليل هذه البيانات يسمح باستكشاف ارتباطات بين مجموعة من البيانات المستقلة لكشف جوانب عديدة ،ومنها على سبيل المثال التنبؤ , كما توفر هذه التنبؤات لصانع القرار أدوات مبتكرة لفهم أفضل للظروف والمعطيات وبالتالى اتخاذ قرارات صحيحة تحقق الأهداف المطلوبة.

من هنا ظهر ما يسمى التنقيب عن البيانات كتقنية تهدف إلى استنتاج المعرفة من كميات هائلة من البيانات،تعتمد على الخوارزميات الرياضية والتي تعتبر أساس التنقيب عن البيانات .

تقدُم مجال الذكاء الاصطناعى أصبح مذهلًا، خصوصًا بعد إسهامه في تحليل البيانات آليًا. فلم تعد الأنظمة تبرمج سطرًا بسطر بل أصبح بإمكانها التعلم وتطوير نفسها.