

ملتقى دور المعلومات في الانتقال نحو اقتصاد المعرفة



BIG DATA

Dr. Basma BOUKENZE
Université Hassan Premier ,SETTAT

Plan



- C'est quoi ce Big Data
- Enjeux du Big Data
- Impact du Big Data

Introduction

L'omniprésence des nouvelles technologies a rendu les tâches quotidiennes faciles et simples, En revanche, elle a donné naissance à une croissance exponentielle du volume de données disponibles, à tel point que les outils traditionnels de gestion des données deviennent incapables de les gérer.

Cette inondation de données, baptisée dans la communauté scientifique par le Big Data, est issu de plusieurs sources, et a touché tous les domaines et secteurs.

le Big Data est le terme phare de ses dernières années dans les entreprises :constitue une aide à la prise de décision au même titre que la business intelligence puisqu'il permet de **mieux cerner les besoins des clients** ou même d'**anticiper leurs futures consommations**.

Introduction au Big Data

C'est quoi Big Data?

Qu'est ce qui rends la data, "Big" Data?

Définition du Big Data

- No single standard definition...

“**Big Data**” , littéralement les grosses données, parfois appelées données massives, est une expression anglophone utilisée pour désigner des ensembles de données qui deviennent tellement volumineux **qu'il en devient difficiles à travailler avec des outils classiques de gestion de base de données ou de gestion de l'information**. L'on parle aussi de datamasse en français par similitude avec la biomasse.

هي بيانات تولدت من خلال استخدامنا المتزايد للأجهزة الرقمية، والأدوات والمنصات المدعومة من شبكة الإنترنت في حياتنا اليومية. وفي أي لحظة معينة، يستخدم مئات الملايين من الأفراد في أنحاء العالم بعضًا من الهواتف المحمولة في العالم –البالغ عددها من سبعة مليارات إلى ثمانية مليارات هاتف– لإجراء مكالمات، أو لإرسال رسائل نصية أو بريد إلكتروني.

كل هذه الأنشطة يترك أثرًا رقميًا. وتُشكّل هذه المعلومات الرقمية في مجموعها الجزء الأكبر من البيانات الضخمة.

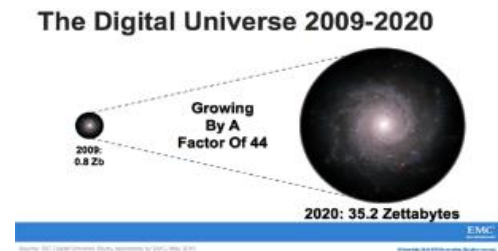
ومنذ عام 2012، يجري إنتاج أكثر من 1.2 زيتابايت من البيانات سنويًا؛ أو ما يكفي لملء 80 مليار جهاز 'آي فون' ساعة 16 جيجابايت (التي قد تُشكّل على امتدادها أكثر من 100 دورة حول الأرض).

إن حجم هذه البيانات في نمو سريع لذلك فإن الحجم، والسرعة، والتنوع هي 'الخصائص' الثلاث التي تميز البيانات الضخمة، وغالبًا ما يُضاف إليها القيمة التي يمكن استخراجها منها باعتبارها الخاصية الرابعة.

Caractéristiques du Big Data:

1- (Volume)

- **Volume**
- les masses de données à traiter sont sans cesse croissantes.
 - 44x increase from 2009 2020
 - From 0.8 zettabytes to 35zb

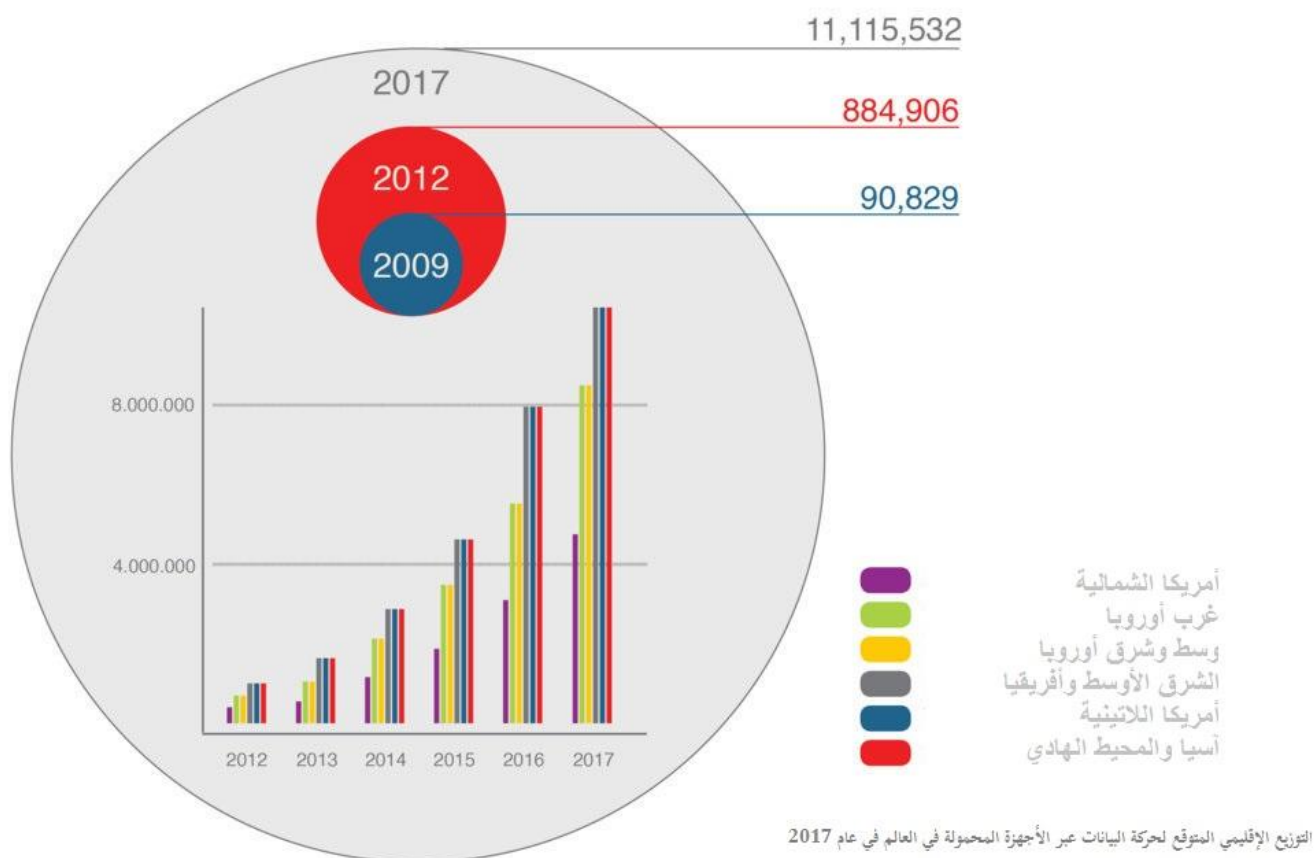


C'est le V le plus associé au Big Data.

Ex: Facebook stocke des photos. Cette déclaration n'a rien d'impressionnant... jusqu'à ce que vous réalisiez que Facebook compte plus d'utilisateurs que la Chine n'a d'habitants. Chacun de ces utilisateurs y stocke de nombreuses photos.

Facebook stocke ainsi environ 250 milliards d'images.

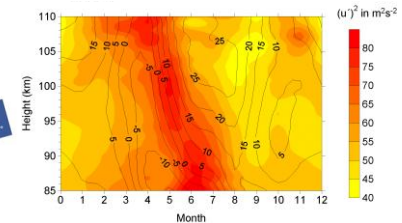
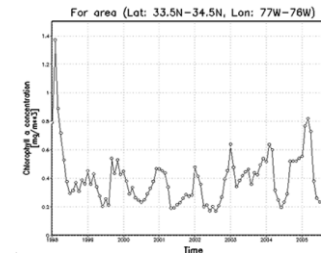
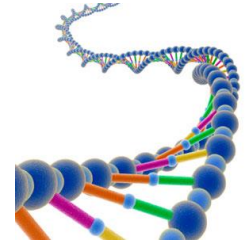
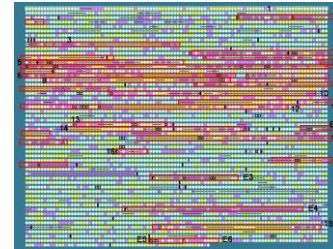
بيانات الأجهزة المحمولة في العالم – نمو الحركة وتوقعاتها (تيرابايت في الشهر)



Caractéristiques du Big Data:

2-Complexité (Variété)

- Variété car les données sont de formats très variés et pas toujours structurées
- Text, numerical, images, audio, video, time series, social media data, etc...



Caractéristiques du Big Data:

3-Vitesse (Vélocité)

La dimension vélocité correspond à la vitesse à laquelle doivent être collectées, analysées et exploitées les données Big Data. De plus en plus souvent, ces données sont à traiter en temps réel.



Facebook → 10 TB par jour

Twitter → 7 TB par jour

Volume

Gigabyte

Térabyte

Pétabyte

Exabyte

Vélocité

Génération

Capture

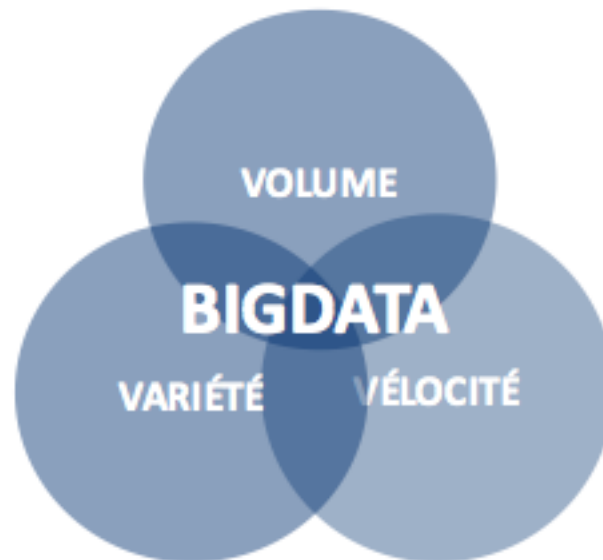
Partage

Variété

Texte

Images

Vidéos



أمثلة من الواقع على البيانات الضخمة:

لدى واتس آب أكثر من مليار مستخدم، و يتم تداول أكثر من 42 مليار رسالة و حوالي 1.6 مليار صورة بشكل يومي.

فيسبوك تتعامل مع أكثر من 50 مليار صورة من مستخدميها.

IBM

إننا ننتج 2.5 كوينتيليون بايت من البيانات كل يوم (الكوينتيليون هو الرقم واحد متبوعاً بـ 18 صفراً). هذه البيانات تنبع من كل مكان، مثل المعلومات حول المناخ والتعليقات المنشورة على مواقع التواصل الاجتماعي والصور الرقمية والفيديوهات ومعاملات البيع والشراء وجوجل تتعامل مع حوالي 100 مليار عملية بحث في الشهر.

Deux autres dimensions se sont ajoutées par la suite à ces 3 dimensions initiales des données Big Data (on parle aussi des « 5 V du Big Data ») :

✓ Véracité ()

✓ Valeur()

✓ Véracité (مصادقية)

La véracité (ou validité) des données correspond à la confiance que ces Big Data inspirent aux décideurs. Si les utilisateurs de ces données doutent de leur qualité ou de leur pertinence, il devient difficile d'y investir davantage.

صدق (أو صحة) البيانات أن هذه البيانات الكبيرة مصدر إلهام في صناعة القرار.

✓ Valeur (قيمة)

Cette dimension s'attache à un aspect étonnant du Big Data. Une donnée prise isolément peut ne pas avoir grand intérêt, mais prise dans un ensemble plus large, elle prend tout son sens.

هذا هو الحال، على سبيل المثال، تويت تتعلق بحدث إخباري. قد لا يكون محتوى تويت ذا أهمية، ولكن إذا تكررت وعلق عليها آلاف المرات، فإن عدد الرسائل المرتبطة بهذه التغريدة التي تصبح معلومات قيمة (رأي مشترك، موضوع هو الإجماع أو الجدل ...).

Générateurs du Big Data



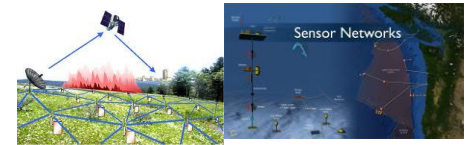
Social media and networks
(all of us are generating data)



Scientific instruments
(collecting all sorts of data)



Mobile devices
(tracking all objects all the time)



Sensor technology and networks
(measuring all kinds of data)

قامت اللجنة الاقتصادية لأوروبا، بتقديم تقرير بعنوان ”ماذا تعنيه البيانات الضخمة للإحصاءات الرسمية“، و قد أوردت فيه تصنيفا لمصادر البيانات الضخمة على النحو التالي:

المصادر الناشئة عن إدارة أحد البرامج، سواء كان برنامجا حكوميا أو غير حكومي، كالسجلات الطبية الإلكترونية وزيارات المستشفيات وسجلات التأمين والسجلات المصرفية

المصادر التجارية أو ذات الصلة بالمعاملات، الناشئة عن معاملات بين كيانيين، على سبيل المثال معاملات البطاقات الائتمانية والمعاملات التي تجرى عن طريق الإنترنت بوسائل منها الأجهزة المحمولة.

مصادر شبكات أجهزة الاستشعار، على سبيل المثال، التصوير بالأقمار الصناعية، وأجهزة استشعار الطرق، وأجهزة استشعار المناخ.

مصادر أجهزة التتبع، على سبيل المثال تتبع البيانات المستمدة من الهواتف المحمولة والنظام العالمي لتحديد المواقع.

مصادر البيانات السلوكية، على سبيل المثال، مرات البحث على الإنترنت عن منتج أو خدمة ما أو أي نوع آخر من المعلومات، و مرات مشاهدة إحدى الصفحات على الإنترنت.

مصادر البيانات المتعلقة بالآراء، على سبيل المثال، التعليقات على وسائل التواصل الاجتماعي.

The Model Has Changed...

- **The Model of Generating/Consuming Data has Changed**

Old Model: Few companies are generating data, all others are consuming data



New Model: all of us are generating data, and all of us are consuming data



Analyse Prédictive

التحليل التنبئي

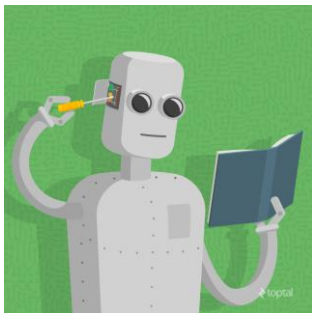
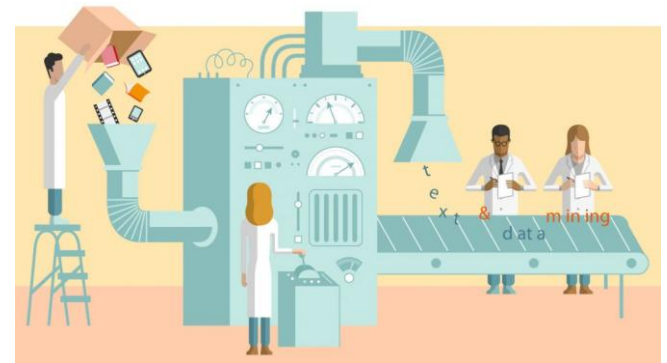
Les analyses prédictives sont le résultat pratique du Big Data et de la Business Intelligence (BI). Elles permettent d'exploiter les immenses quantités de données collectées par de nombreuses entreprises auprès de leurs clients, de leurs marchés, des réseaux sociaux, des applications en temps réel, ou encore du Cloud. En dégagant des informations tangibles.

يعتبر التحليل التوقعي أو التحليل التنبئي نوعاً من التنقيب في البيانات هو مجال للتحليل الإحصائي للبيانات هدفه استخلاص معلومات حول التغيرات السلوكية المستقبلية , يتأسس التحليل التوقعي على فهم العلاقة بين المتغيرات التي تتسبب في الأحداث و المتغيرات المتغيرة مع الأحداث أو المتوقعة (المظاهر و الأسباب) المنبثقة عن تجارب في الماضي و استعمال هذه العلاقات من أجل توقع المستقبل . تجدر الإشارة إلى أن نتائج التحليل التوقعي تتأثر بشكل كبير بجودة الفرضيات و مستوى تحليل البيانات

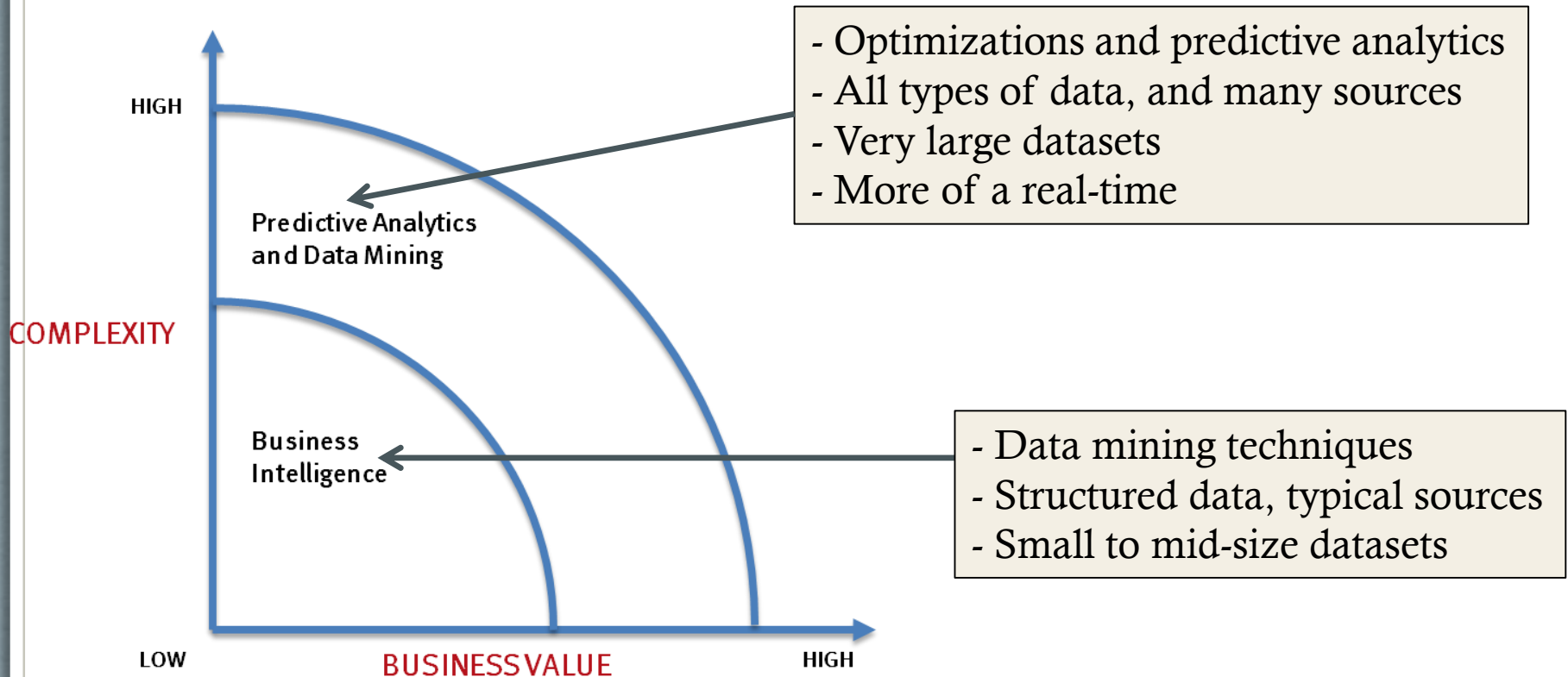


- Le terme analyse prédictive rassemble de nombreuses technologies d'analyse de données et autres techniques:

- Datamining
- Machine learning



What's driving Big Data



What Technology Do We Have For Big Data ??

يقول ألين ميتشل، مدير تقني أول لمنطقة الشرق الأوسط وشمال إفريقيا في الشركة المتخصصة في إدارة البيانات "كومفولت": "تتطلب البيانات الكبيرة مجموعة كبيرة من المهارات الخاصة بالتحليل والبرمجة لا تحتاجها أدوات التحليل الموجودة اليوم".

يضيف: "تتضمن تلك المهارات خلفيات قوية في الرياضيات والتحليل الإحصائي بالإضافة إلى معرفة واسعة بأحدث لغات البرمجة الإحصائية وتقنيات وضع نماذج تحليل متخصصة".

The Big Data Landscape

Apps

Vertical Apps



Operational Intelligence



Data As A Service



Ad / Media Apps



Business Intelligence



Analytics And Visualization



Infrastructure

Analytics Infrastructure



Operational Infrastructure



Infrastructure As A Service



Structured Databases



Technologies



Acquisition

Marshalling

Analysis

Action

Data Acquisition



VLDW and BI Appliances



Analytics



BPM & Action



Including Complex Event Processing (CEP) tools

Data Providers



No SQL



Data Virtualization COMPOSITE SOFTWARE



Content Management



BI Tools



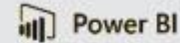
Capgemini - Capping IT off
Manuel Sevilla - 2012

Data Governance

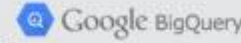


Big Data Technology

Visualization & Analytics



Compute



Storage



Distributions & Data Warehouse



The Big Data technology stack is evolving rapidly

AMAZON

تتوقع ما قد يعجبك من منتجاتها في كل مرة تتسوق فيها

NETFLIX

تتوقع أي البرامج والعروض التي تفضل مشاهدتها.

GOOGLE

كذلك تتنبأ بالطريقة التي ترد فيها على الرسائل الواردة إلى بريدك الإلكتروني.
وهذا كله من خلال تقنيات تحليل البيانات التي ما زال الخبراء يعملون على
الاستفادة منها لأغراض أخرى، منها:

opportunités du Big Data

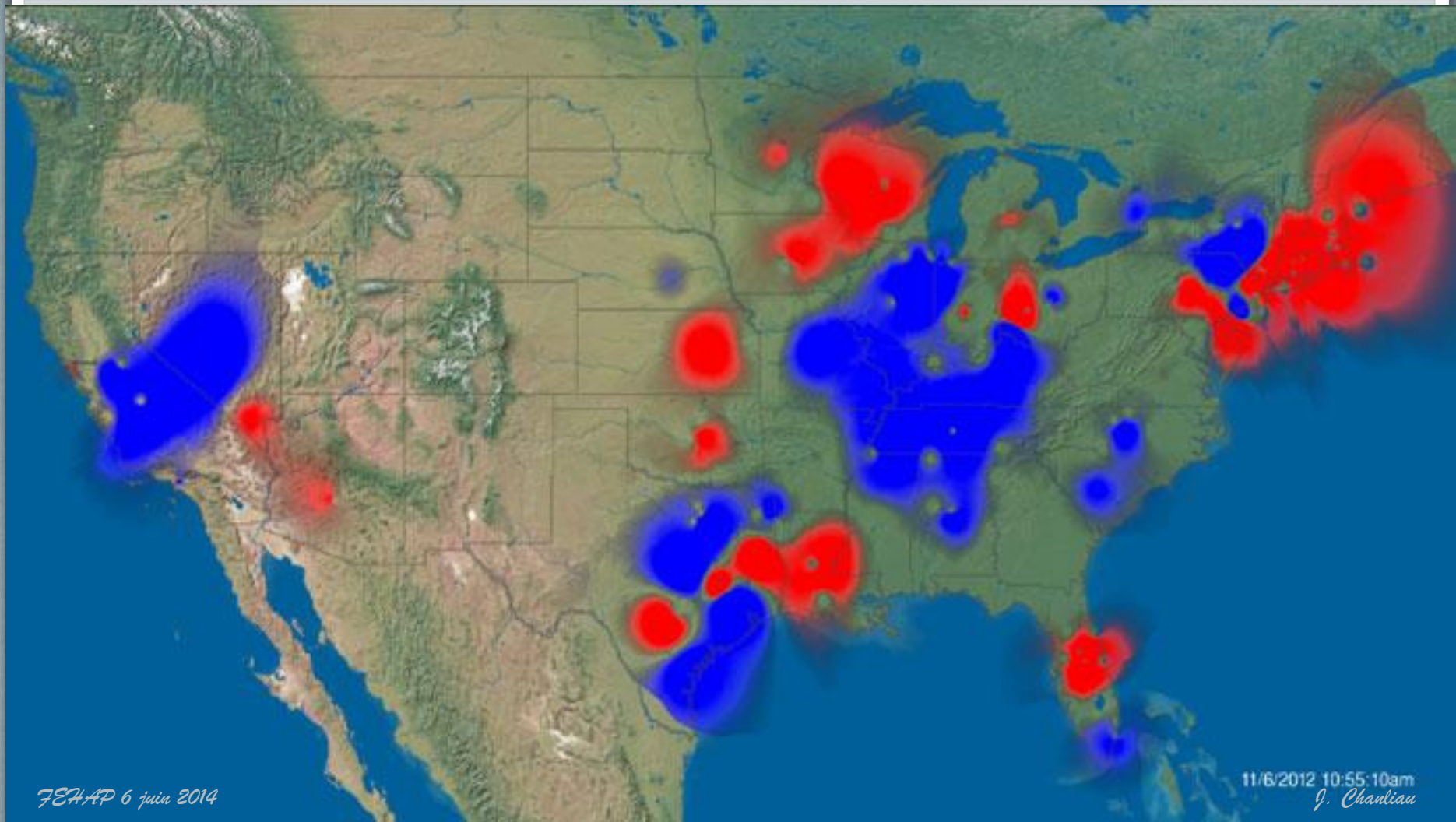
التسويق → معرفة العملاء التي تمكنها البيانات الكبيرة بحيث يصبح من الممكن ,
للشركة إضفاء الطابع الشخصي وتقسيم العروض الترويجية بدقة لا مثيل لها

التجارة → يمكن استخدام استطلاعات الرأي للعملاء استنادا إلى تفاصيل الطلب وتواتر
المشتريات لتحسين وضع المنتج على الرفوف أو نقاط البيع وتحديد الأسعار.

الإنتاج ← يمكن استخدام تحليل البيانات لتحسين العمليات، وتوفير الطاقة والموارد
الطبيعية، وتحسين إدارة المخزون

المالية ← البنوك وشركات التأمين تستخدم البيانات كبيرة لتحديد مستوى خطر العميل
المتعاقد وتسليط الضوء على السلوكيات المشبوهة أو كشف الاحتيال المحتمل.

En 2012, l'analyse des tweets des américains au sortir des bureaux de vote avait permis de prédire les résultats de l'élection présidentielle.



Santé → Le big data rend théoriquement possible une **médecine préventive et personnalisée** en lien avec des appareils connectés mesurant les données biométriques des patients et visant à leur **proposer des conseils ou des traitements** appropriés.

In Search of Solutions



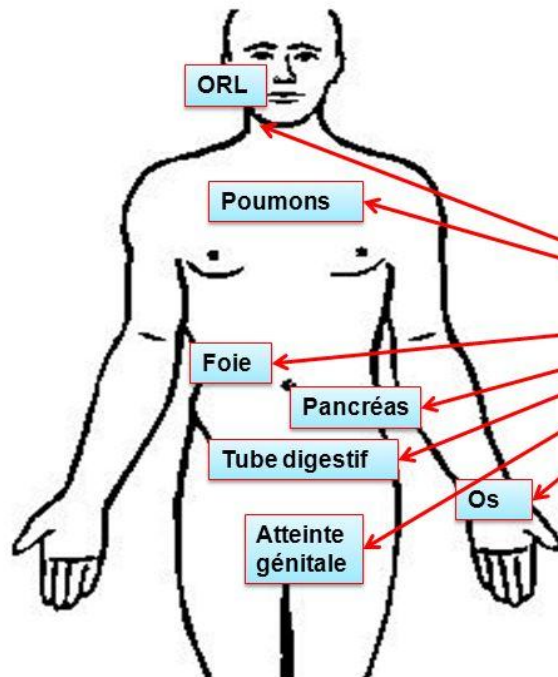
Do You Have a Big Data Problem?

Gartner

Médecine individualisée/Prédictive

La manière de se soigner en 2025 sera très différente de ce que nous connaissons. Ainsi, il est d'ores et déjà acquis que la médecine saura personnaliser son diagnostic en fonction des caractéristiques propres de chaque individu et notamment de son génome. Elle pourra adapter les modalités d'interventions médicamenteuses ou chirurgicales aux spécificités de chaque patient,

La médecine personnalisée connaît actuellement un essor important dans le domaine de la mucoviscidose



*Médecine dite
Personnalisée ou
individualisée*
Les thérapies ciblées

*Médecine Prédicative &
Préventive*
*Les gènes modulant la
sévérité des différentes
atteintes*

Médecine Participative

Médecine des « 4P »

- ✓ Prévoir les pathologies
- ✓ Anticiper le traitement
- ✓ Réduire les couts liée au traitement
- ✓ diminuant ainsi le coût des soins à l'échelle d'une société. L'individu et ses caractéristiques propres seront, plus que jamais, au cœur de la médecine de demain avec une forte diminution des risques associés aux soins.

يقول ريغاناثان كريشنا مورثي، محلل أول في مؤسسة الأبحاث فروست آند سوليفان: "تزداد معدلات استخدام البيانات الكبيرة في منطقة الشرق الأوسط في الوقت الذي يدرك فيه العديد من المدراء التنفيذيين أهمية دورها في أعمالهم".

وبالنظر إلى ذلك، تشير التقديرات إلى أن 80% من بيانات الشركات غير منظمة وهناك حاجة كبيرة إلى إيجاد حل لذلك

ويضيف: "بلدان مثل السعودية والإمارات تأتي في المرتبة الأولى من حيث مستويات الاستخدام ومع ذلك فإن 25% من المؤسسات لم تدرك بعد الأهمية الكبيرة للبيانات الكبيرة كأداة. وحتى تلك التي تدرك أهمية الانتقال للعمل بالبيانات الكبيرة فإنها لم تقم بذلك خشية عدم القدرة على معرفة حجم عائدات الاستثمار.

Conclusion

- Submergés par de gros volume de données ,
- Ce Big data constitue une vraie richesse ,
- À la bonne prise de décision et au moment opportun,
- Toutes les entreprises doivent faire face à ce déluge de données et d'en tirer le maximum de profit
- Gains
- développement économique



شكرا



لكم خالص
الشكر
والتقدير