

# البيانات الضخمة في المكتبات : بين التعقيد والتبسيط

د. أحمد سعد الدين بسيوني  
مكتبة الإسكندرية - مصر

# أهمية العرض



- أحد أهم التقنيات الواجب توافرها في المكتبات لمواكبة مجتمع المعرفة .
- التقنيات المستخدمة في المكتبات حالياً لا تدعم تحليل البيانات الضخمة .
- يعد من الموضوعات الحديثة نسبياً في أدبيات المكتبات مع ندرة الدراسات العربية .
- مع التقدم التقني وخفض تكلفة الأجهزة والبرمجيات، ومن ثم انتشار تحليل البيانات الضخمة بين مختلف الأنشطة التجارية والأكاديمية وغيرها، بعدما كان قاصراً على المؤسسات الكبيرة والحكومات

# الإشكالية



كيفية الاستفادة من هذا الكم الهائل من البيانات

والتي اغلبها غير منتظمة

وسريعة النمو ومتنوعة الأشكال

# ماهية البيانات الضخمة



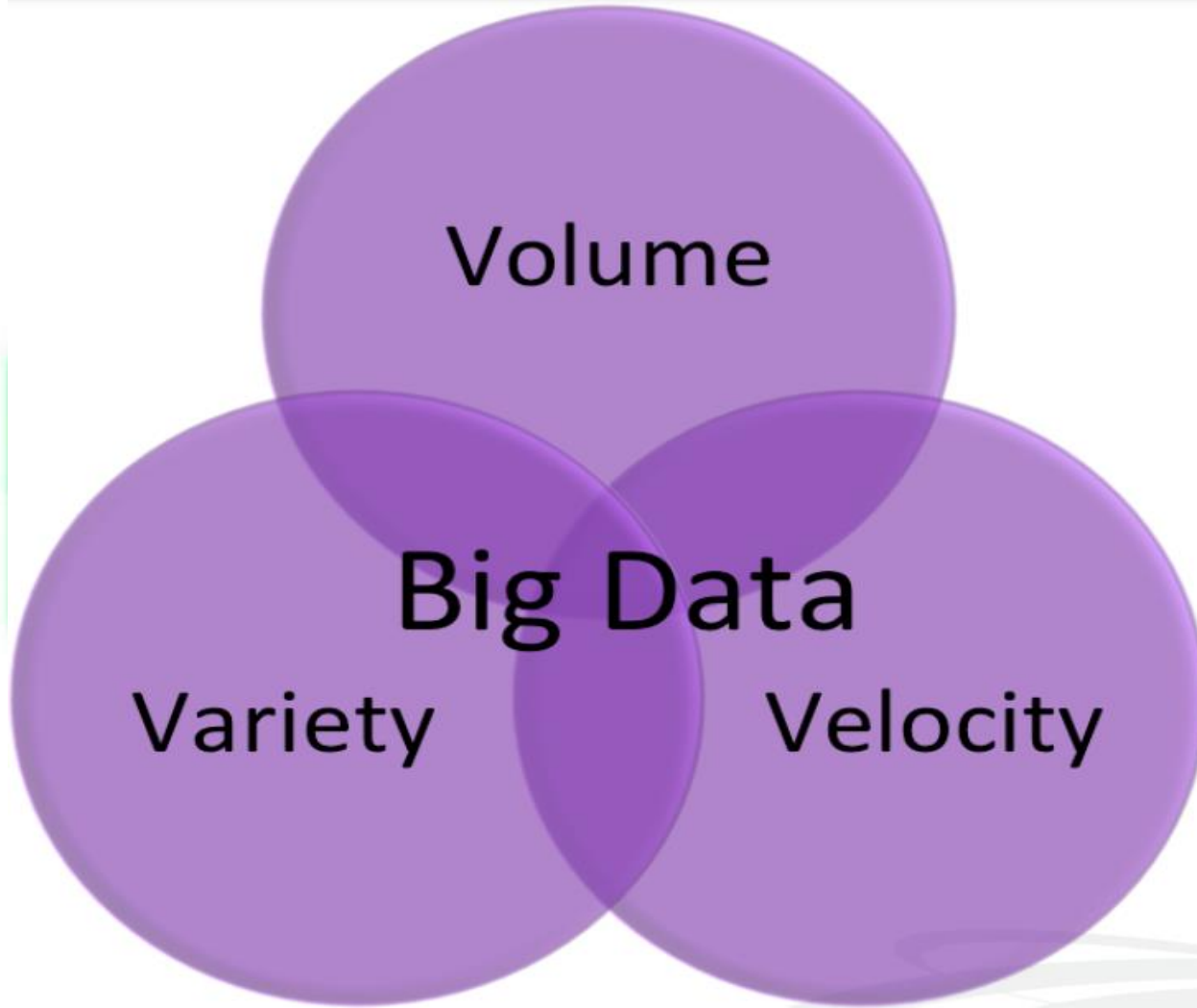
هي ذلك الكم الهائل من البيانات المتداولة على الانترنت، سريعة الإنتاج والانتشار، مختلفة المصادر، متنوعة الأشكال، مع ما تحمله من قيمة.

## أهميتها

- وفرتها المفردة مع ما تحمله من قيمة .
- يكشف تحليل البيانات الضخمة عن الاتجاهات الاجرامية أو التجارية أو السلوكية أو الوظيفية لدى أصحاب هذه البيانات .
- تستخدم البيانات في مكافحة الإرهاب والتخطيط ورسم السياسات وتسهيل الإجراءات والتسويق والإتاحة الفورية للمستفيدين .



## خصائصها



variety	التنوع
velocity	السرعة
volume	الحجم
veracity	المصداقية
value	القيمة
visualization	التصور
viscosity	الالتصاق
virality	الحيوية

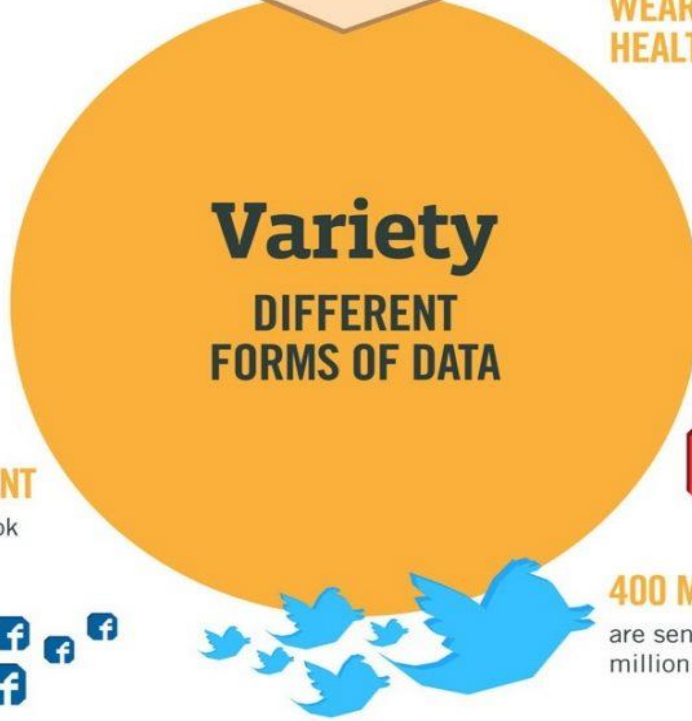
# Variety التنوع

As of 2011, the global size of data in healthcare was estimated to be

**150 EXABYTES**  
[ 161 BILLION GIGABYTES ]



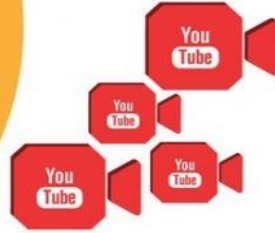
**30 BILLION  
PIECES OF CONTENT**  
are shared on Facebook  
every month



By 2014, it's anticipated there will be

**420 MILLION  
WEARABLE, WIRELESS  
HEALTH MONITORS**

**4 BILLION+  
HOURS OF VIDEO**  
are watched on  
YouTube each month



**400 MILLION TWEETS**  
are sent per day by about 200  
million monthly active users



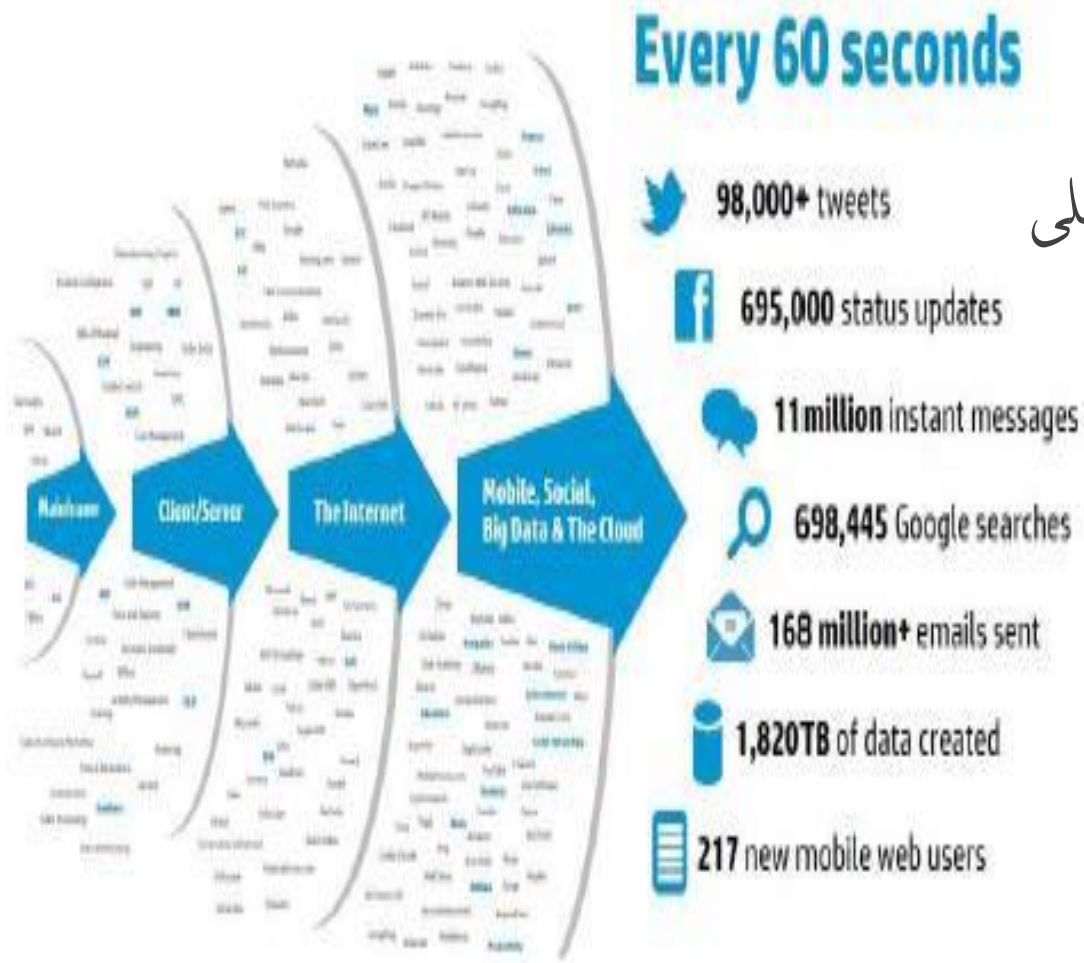
البيانات تغيرت من مجرد أرقام ونصوص تقليدية فقط. أصبحت تضم:

- مقاطع الفيديو.
- صور فوتوغرافية.
- تسجيلات الصوتية.
- دردشات على الماسنجر.
- رسائل على الوتساب.
- تداولات أسهم البورصة.
- بيانات حركة المرور.
- مشاركات على الفيسبوك.
- تغريدات على تويتر.
- رسائل البريد الإلكتروني.
- مؤشرات وأخبار الطقس.
- محددات المواقع الجغرافية.

تعليقات على منتجات أو خدمات أو منشورات. الخ.

وتنوع البيانات ما بين بيانات منظمة أو مهيكلة وتمثل النسبة الأقل، وبيانات غير منظمة أو غير مهيكلة وتمثل الجزء الأكبر، أو خليط بين الاثنين تسمى بالبيانات شبه المنظمة أو شبه المهيكلة.

# Velocity السرعة



تعنى السرعة المذهلة التي يتم بها إنشاء البيانات ومعالجتها ومشاركتها على مواقع التواصل الاجتماعي . وبلايين المستخدمين يمارسون أنشطة متواصلة على كل من مواقع التجارة الإلكترونية، المدونات، الشبكات الاجتماعية، والتدفقات الحظية المستمرة للبيانات لاستبيان ميول المستهلكين الآنية ويبقى التحدي في : كيفية التفاعل مع طوفان المعلومات في الوقت المناسب ؟

# المجموع Volume



• وصلت بيانات العالم الرقمي اليوم إلى نحو 44 ريانايت.

• I.3 ترليون كيلومليقبحث سنوي على KB ووجل فقطريبا IO30

• فيسبوك يتداول 3I.3 مليون رسالتة تقريبا كل دقيقة. MIB تقريباً IO60

• ويستعرض 2.77 مليون مقطع فيديو. GB = تقريباً IO90  
• يتم تحميل نحو 300 ساعة فيديو من يوتيوب كل دقيقة.

• لدينا نحو I. TB بايليتار مستخدم للهواتف الذكية تقريباً IOI2

• ينتج تويتر أكثر من 100 تيرابايت من الـ PBات كل يوم تقريباً IOI5

• وينتج فيسبوك نحو IO تيرابايت من البيانات كل يوم تقريباً IOI8  
• إكسبايت = تقريباً EB

• ZB = تقريباً IO2I زيبايت

# هل تعد بيانات المكتبات بيانات ضخمة؟

الحجم: المواد والتعاملات المرقمنة بالمكتبة تتراكم لتصبح بيانات ضخمة.

التنوع: كتب ورسائل جامعية مرقمنة، دوريات الكترونية، أفلام ومقاطع مصورة، تسجيلات صوتية، صور فوتوغرافية.  
السرعة: تحرك للملفات بين المكتبات، الأبحاث الجارية، مجموعات ومقتنيات المكتبة المتزايدة، التعاون بين المكتبات، الإحاطة الجارية، البث الانتقائي. الاستفسارات الجارية.



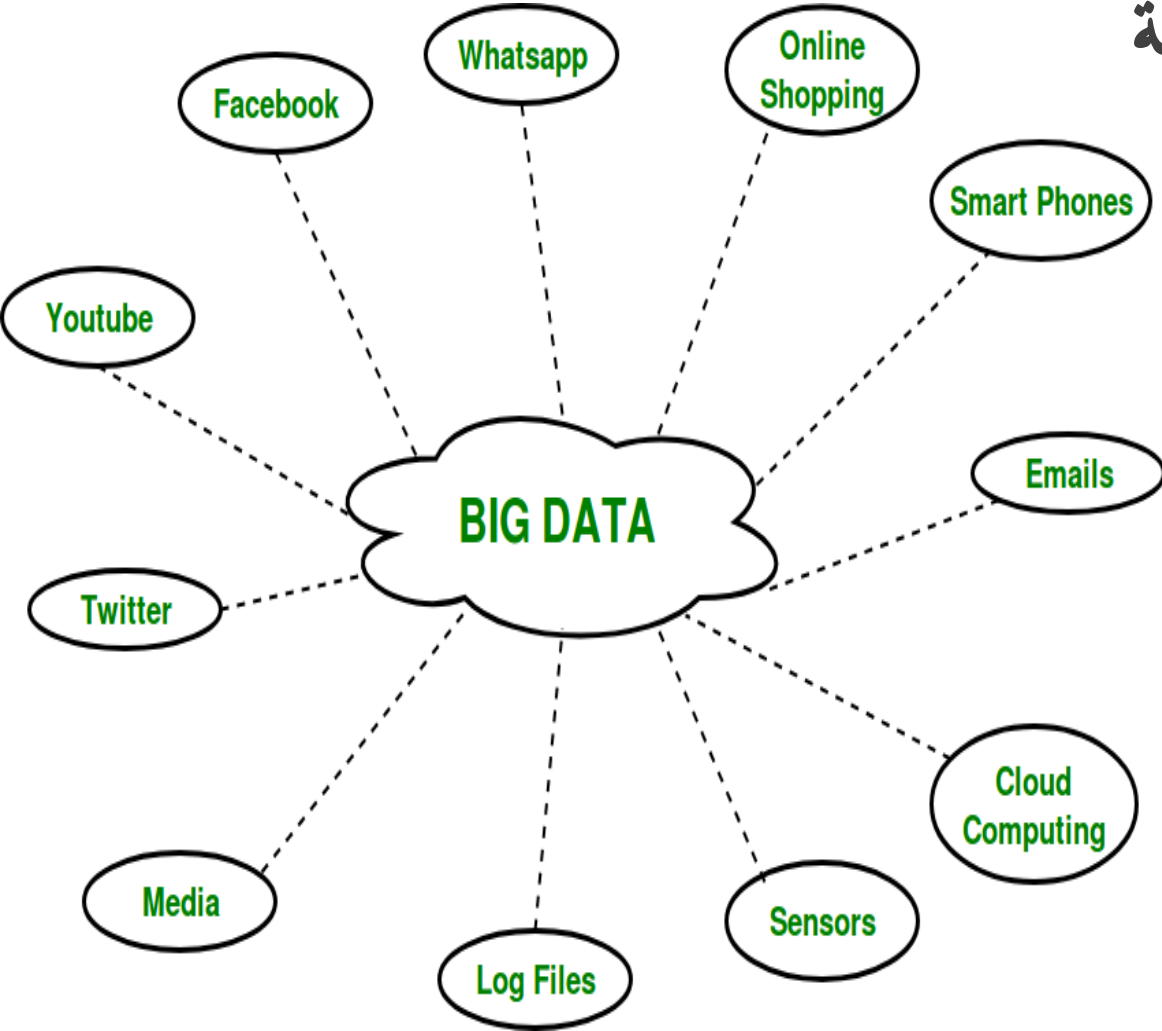
## مصادر البيانات الضخمة بالمكتبة

مصدرين أساسيين .

أولاً: وسائل التواصل الاجتماعي الإلكترونية بين المكتبة والمستفيدين .  
ثانياً: قواعد البيانات الإلكترونية توفر نحو 30% من بيانات المكتبة .  
(سجلات إعاره، استفسارات بحثية، ملفات المستفيد، . . . الخ.)

مصادر ثانوية أخرى .

أهمها الأنشطة الثقافية والفنية (المؤتمرات، والندوات، والمحاضرات،  
وورش العمل، والمهرجانات، والمعارض الفنية... الخ.)







## تقنيات البيانات الضخمة

هو برنامج أو منصة يمكن من خلالها استخراج القيمة المطلوبة من البيانات الضخمة والعمل عليها .

وهو يعمل على الربط بين عدد من الخوادم والتعامل معها كأنها سعة تخزينية واحدة، وبالتالي تمكن برامج

تحليل البيانات من التعامل معها والاستفادة منها .



أولاً : أدوات تحليل الدفقات حيث تتم تصفية البيانات أولاً ثم تحليلها، ومن أشهرها :

Apache Hadoop, Google MapReduce, Microsoft Dryad, Apache Mahout.



ثانياً : أدوات تحلل دفق البيانات في الوقت الحاضر باعتبارها قابلة لاستخلاص النتائج، ومن أشهرها:

Apache Storm, Apache S4, Project Spark, MOA.



ثالثاً : أدوات التحليل التفاعلي الذي يسمح للمستخدمين بمعالجة البيانات وفق تحليلهم الخاص، ومن أشهرها :

Apache Drill, D3.

# أوجه الاستفادة من البيانات الضخمة في المكتبات



أولا : تعزيز ودعم أو إعادة الهيكلة الإدارية .

ثانيا : إدارة وتسويق الخدمات والأنشطة .

ثالثا : خفض تكلفة ووقت التشغيل .

رابعا : قواعد بيانات المستفيدين .

سادسا : دراسات المستفيدين .

خامسا : تنمية الموارد البشرية .

سابعا : إدارة مجموعات المكتبة .

# الرؤية الإستشرافية للإفادة من البيانات الضخمة في المكتبات



- اكتشاف الفرص غير المستغلة ومواقع الأهدار ومن ثم توفير المزيد من الجهد والوقت والمال.
- رفع الكفاءة والفاعلية للاستغلال الأمثل للموارد المتاحة.
- دعم القدرة على التنبؤ بالمستقبل للاستعداد المبكر له.
- تحسين العمليات وتقديم خدمات وأنشطة أفضل.
- إتخاذ قرارات مبنية على واقع، مما يرفع درجة صواب القرارات.

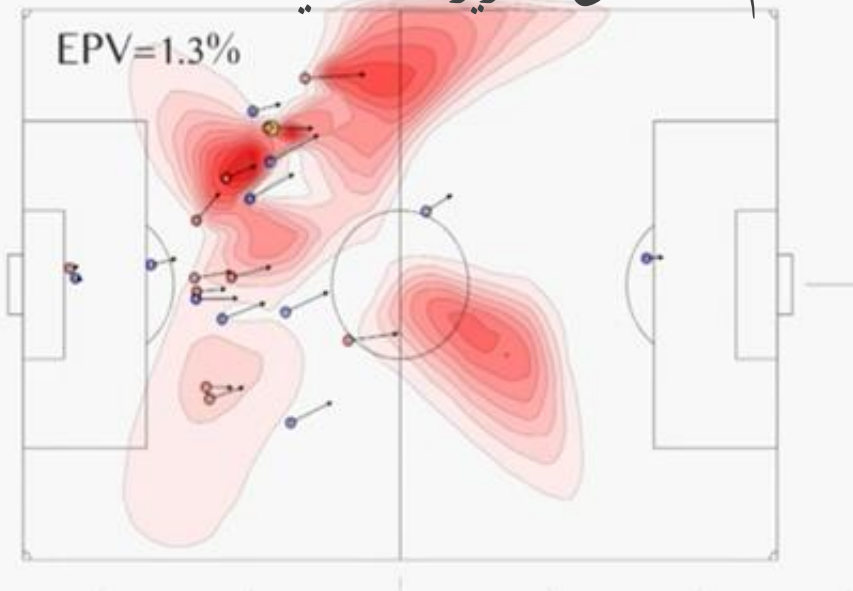
# نموذج الإفادة من البيانات الضخمة في كرة القدم



قامت إدارة البحوث في نادي ليفربول بتحليل البيانات الضخمة لفريقيها وللفرق المنافسة.

تم جمع المعلومات من الفيديوهات وتحليلها والخروج بإحصاءات عن تحركات اللاعبين والثغرات التي تحدث داخل

الملعب، واستلام الكرة، وفرص واوقات بناء الهجمة، وأماكن المساحات الفارغة للخصم، ومناطق التمير والتسديد.



والخروج باحتمالات لسد الثغرات وتسجيل الأهداف.

تم وضع برامج التدريب وفقا لهذه النتائج.

# نموذج الاستفادة من البيانات الضخمة في كرة القدم

Will Spearman  
"بيانات التتبع"  
"تحليل النقاط المتحركة"

ونتيجة لهذه استطاع ليفربول أن ينتقل من  
الثامن :2016-2015

...

...

...

الأول

:2020-2019

...

...

الثالث

2024-2023

فإذا كانت فرق كرة القدم قد استغلت تحليل مجموعات البيانات الضخمة لتطوير نتائجها.  
فواجب علينا أن نقوم بهذا الدور لتطوير وتحسين أنشطة وخدمات المكتبات.  
ولاسيما وأن مجال عملنا هو توفير وإتاحة البيانات والمعلومات.

## التحديات التي تواجه المكتبات

قدرات  
واحتياجات  
وتوقعات  
المستخدمين

البيانات: الحجم،  
الفورية، والتنوع

المهارات الفنية  
لدى أمناء  
المكتبات

المالية: تكاليف  
البرامج والأجهزة  
والتدريب.

الإدارية:  
خصوصية  
وسرية وأمن  
البيانات،

المعالجة:  
المفاضلة،  
التحليل

# استشراف المستقبل



رقمنة  
المشاعر



# التوصيات :



1. عقد الندوات والبرامج للتوعية بأهمية البيانات الضخمة بين المكتبيين.
2. وضع البرامج التدريبية للمكتبيين للتعامل مع البيانات الضخمة (تقنيا وإداريا وتنظيميا).
3. تضمين البيانات الضخمة في المقررات الدراسية.
4. تشجيع باحثي الدراسات العليا على القيام بدراساتهم حولها.
5. إنشاء البنى الأساسية وإعداد الخطط الاستراتيجية الخاصة بها.
6. عقد وتعزيز شراكات التعاون بين المكتبات في البيانات الضخمة مع مشاركة الخبرات.

وأخيراً:

جمع وتحليل البيانات الضخمة وسيلة وليس هدف . . . المهم:

( كيفية التوظيف الجيد للمعلومات المستخرجة منها )

( وهذا لا يعني عن الإدارة )

فهم من يقودون المركبة، وما مخرجات البيانات الضخمة إلا إمكانيات  
ومؤشرات تحتاج إلى قرارات التوجيه من الإدارة.

