Dirassat & Abhath
The Arabic Journal of Human and Social Sciences

مجلة دراسات وابحاث
المجلة العربية في العلوم الإنسانية والاجتماعية

EISSN: 2253-0363
ISSN : 1112-9751

إدارة البيانات الضخمة (Big Data) في الشركات التقنية: دراسة حالة شركتي جوجل Googleوميتا - Meta

Big Data Management in Technology Companys: Case Study of Google and Meta

قوت سهام kout sihem
أستاذ محاضر أ، جامعة عبد الحميد مهري قسنطينة2، معهد علم المكتبات والتوثيق، مخبر اتكنولوجيا المعلومات ودورها في التنمية الوطنية
Maitre conférence classe A ,university of Abdelhamide Mehri constantine2, Laboratory New Technologies and Their Role in National Development

sihem.kout@univ-constantine2.dz

المؤلف المرسل: قوت سهام sihem.kout@univ-constantine2.dz

تاريخ الاستلام: 2022-07-24 تاريخ القبول : 2022-12-23

الملخص باللغة العربية

هدفت هذه الدراسة إلى التعريف بموضوع إدارة البيانات الضخمة، في ظل تطور تقنيات التخزين والتكنولوجيا الجديدة، التي أصبحت منتجة للبيانات المنظمة وغير المنظمة، كالحواسيب والهواتف النقالة، وهناك العديد من المصادر الأخرى التي تولد كماً هائلاً من البيانات على غرار مواقع التواصل الاجتماعي، ما جعل العديد من الشركات التقنية الكبرى والمتميزة في العالم تفكر في كيفية التعامل مع هذه البيانات، واستغلالها في مجال عملها عن طريق مجموعة من البرامج والتقنيات للتعامل مع هذه البيانات الضخمة. وسنستعين في بحثنا بشركتي جوجل Google وميتا Meta كمثالين عن الشركات التقنية.

وخلصت الدراسة إلى نتيجة أساسية مفادها أن الهدف الرئيسي من توفير خدمات مجانية متاحة للجميع من قبل شركتي جوجل وميتا هو تحصيل أكبر قدر من البيانات عن ملايين المستخدمين التي تعتبر معلوماتهم رأس مال هذه الشركات لاستغلالها لأغراض غير معلنة.

الكلمات المفتاحية: البيانات، البيانات الضخمة، إدارة البيانات الضخمة، جوجل، ميتا.

Abstract:

This study aimed to present the topic of big data management, as storage techniques and new technologies have evolved, and which have become producers of regulated and unregulated data, like computers and mobile phones. There are many other sources that generate a huge amount of data like social media sites, which has led many of the world's largest and distinguished technical companies to think about how to deal with this data, and exploit it in its field through a range of programs and techniques. In our research, we will use Google and Meta as examples of technology companies.

The study concluded that the main objective of providing free services available to all by Google and Meta was to collect the most data on millions of users whose information was the capital of these companies for undeclared purposes.

Keywords: Data, Big Data, Big Data Management, Google, Meta.

مقدمة:

- تنتج المؤسسات والأفراد والآلات والأجهزة كميات هائلة من البيانات بمعدلات متسارعة، وعلى المستوى العالمي، "اعتبارا من يناير 2020 كان هناك أكثر 1.74 مليار موقع على شبكة الاتصال العالمية، 7 مليار طلب بحث على الجوجل يوميًا في جميع أنحاء العالم على الرغم من أن البعض يقولون إنه قد يصل إلى 10 مليار في اليوم، وهناك أكثر من 600 مليون مدونة في العالم في عام 2020"[1]، ويعزى هذا النمو السريع في إنتاج البيانات إلى انتشار الأجهزة والأنظمة المتصلة بالإنترنت، والتحول من أجهزة البث التناظري الى التقنيات الرقمية، والاستخدام المتصاعد والسريع لوسائل الإعلام الرقمية من قبل المؤسسات والأفراد. كذلك نجد أن حصة إنتاج البيانات غير النظامية عبر وسائل التواصل الاجتماعي ومقاطع الفيديو والصور هي حاليا أكبر من حصة البيانات النظامية. بالإضافة إلى الكميات الهائلة من البيانات التي يتم إنتاجها وتخزينها وإتاحتها عبر الشبكات حاليا، ازدادت القدرة على معالجة البيانات بشكل كبير، حيث يمكن الآن إجراء تحليلات معقدة على مجموعات كبيرة جدا من البيانات، فالبيانات التي كان ينظر إليها سابقا وكأنها غير مهمة، يمكن الآن أن تضيف معلومات غاية في الأهمية عند دمجها مع بيانات من مصادر أخرى. فقد أصبح من الممكن معالجة كميات كبيرة من

مجموعة بيانات متنوعة بسرعة عالية في الوقت الفعلي، وهي الخصائص التي يتضمنها تعريف البيانات الضخمة، بالإضافة إلى كونها ذات قيمة وموثوقية.

على الصعيد العالمي، من المتوقع نمو سوق تكنولوجيا البيانات الضخمة والخدمات المتعلقة بها بمعدل أسرع بحوالي سبع مرات من سوق تكنولوجيا المعلومات والاتصالات ككل. ووفقا لجارتنر، وهي شركة رائدة في مجال أبحاث تكنولوجيا المعلومات والاستشارات، فإن البيانات الضخمة تمر الآن بذروة مرحلة التضخم ويتوقع أن تصل إلى القمة الإنتاجية خلال نحو خمس إلى عشر سنوات، أي بعد بضع سنوات من الحوسبة السحابية.

لذلك أمسى من الصعب العمل مع البيانات الضخمة باستخدام أنظمة إدارة قواعد البيانات العلائقية وحزم المحاكاة التقليدية، حيث يتطلب الأمر بدلاً من ذلك "برامج متوازية واسعة النطاق تعمل على عشرات أو مئات أو حتى آلاف الخوادم".وما يُعتبر "بيانات ضخمة" يختلف باختلاف قدرات المنظمة التي تقوم بإدارة المجموعة، وعلى قدرات التطبيقات التي تستخدم بشكل تقليدي لمعالجة وتحليل مجموعة البيانات في النطاق الخاص بها، فبالنسبة لبعض المنظمات، ربما تؤدي مواجهة مئات الجيجا بايت من البيانات لأول مرة إلى إعادة النظر في خيارات إدارة البيانات. وبالنسبة للبعض الآخر، ربما يستغرق الأمر عشرات أو مئات تيرابايت من البيانات قبل أن يصبح حجم البيانات شأناً مهماً.

كما أن الكم الهائل من البيانات التي يجري إنتاجها وتخزينها والعمل على إتاحتها من مواقع متعددة مصدر قوة رئيسيا إلى مجتمع قائم على المعرفة، فهذه البيانات الضخمة من شأنها، في حال إدارتها على نحو صحيح، أن تسهم إسهاما مؤثرا في التنمية الاقتصادية والاجتماعية المستمرة. في حين تدرك المؤسسات والشركات الفرص والمكاسب التي توفرها البيانات الضخمة لجميع قطاعات المجتمع.

ومن هنا تتضح معالم إشكالية هذه الورقة البحثية في التساؤل التالي: ما واقع إدارة البيانات الضخمة في الشركات التقنية؟ وما الدور الذي تلعبه في نجاحها وتميزها؟

المحور الأول: إدارة البيانات الضخمة Big Data

أولا-ماهية البيانات:

سنتطرق في هذا العنصر إلى:

1-مفهوم البيانات

تعرف البيانات بأنها "عناصر خام لا معنى لها" [2]

وهناك من يعرفها بأنها" حروف، جمل، أرقام، أو رموز غير منظمة وغير مرتبطة بموضوع واحد والتي قد لا يستفاد منها في شكلها الحالي"[3] ومن هنا يمكن القول أن البيانات في حد ذاتها لا يمكن أن تقدم معنى ولا يمكن أن يستفاد منها.

وتعرف البيانات بأنها: "مجموعة من الحقائق أو الوسائل أو الإشارات أو الآراء أو الاتجاهات" [4]فهي بذلك تمثل مجموعة من الحقائق المجردة التي تعبر عن أحداث معينة بهيئة رموز أو حروف أو أرقام أو رسوم بيانية يتم جمعها من مصادر مختلقة، أي أنها المادة الخام اللازمة لإنتاج المعلومات.**2- ال**

2- وحدة قياس البيانات

وحدات قياس البيانات هي الوحدات التي تستخدم لحساب مساحات الذاكرة في الحاسوب، وهي تعبر أساساً عن كمية المعلومات المخزنة وتقاس عادة بالبايت ومضاعفاته (بت، بايت، كيلوبايت..)، ويمكن توضيح ذلك في الجدول رقم (01).

حيث يعتبر بت أصغر وحدة قياس للذاكرة وتتكون من رقم واحد إما صفر أو واحد، أما البايت يتكون عادة من 8 بت، في حين الكيلو بايت تساوي حوالي 1024 بايت. وميجابايت تساوي1024 كيلوبايت، وجيجابايت تساوي1024 كيلوبايت، وتيرابايت تساوي 1024 جيجابايت.

3-أنواع البيانات Data types

تنقسم البيانات الى قسمين أساسيين هما:

1-3 البيانات المهيكلة structred Data

وهي بيانات منظمة في شكل جداول أو قواعد بيانات، وهي على عكس البيانات غير المهيكلة تتبع مخطط قاعدة بيانات محددة مسبقا.

وتكون البيانات المنظمة على شكل حقول مثبتة ومنسقة في سجل أو قواعد البيانات العلائقية، والتي تتكون من جداول قواعد البيانات، في قائمة البيانات العلائقية يتم استخدام الجداول لتنظيم المعلومات في صفوف وأعمدة (المخطط يتكون من جداول وحقول) وتحتوي هذه الجداول على أنواع البيانات مثل الأرقام والتواريخ والعناوين وذلك على نحو سلسلة من الأعمدة، والمخطط يحدد كيفية تخزين البيانات والتحكم في الوصول اليها وسهولة معالجتها، وهذا يكون استنادا الى طريقة المخطط عند الكتابة.[5]

2-3 البيانات غير المهيكلة unstructred Data

وهي بيانات الكترونية لا يمكن تصنيفها كالصور والرسوم لبيانية بسهولة، كما تعتبر المشاركات المكتوبة في مواقع التواصل الاجتماعي والفيديوهات والصور، والمدونات الالكترونية ورسائل البريد الالكتروني، كل هذه البيانات مصدرا غنيا للغاية بالمعلومات التي تنمو بوتيرة متسارعة وغير مسبوقة(...)، وهذا يعني أنه لا يمكن استخدام أدوات تحليل البيانات القديمة، وادارة قواعد البيانات التقليدية مع هذه البيانات لأنها ليست منتظمة وفق الهيكل الذي يتعامل معها كالجداول مثلا، منصات التدوين والتي لها العديد من العناصر المنظمة منها التاريخ، ووقت النشر والمحتوى، والروابط التشعبية وما الى ذلك، ما يجعل من البحث والتحليل أكثر صعوبة بكثير من البيانات المهيكلة.[6] ويمكن توضيح البيانات المهيكلة وغير المهيكلة في الشكل رقم (02).

ثانيا- ماهية البيانات الضخمة Big Data

يمكن عرض ماهية البيانات الضخمة من خلال العناصر التالية:

1- مفهوم البيانات الضخمة

البيانات الضخمة مصطلح يستخدم لوصف مجموعات بيانات كبيرة للغاية أو على درجة من التعقيد أو التي تتطلب قدرا كبيرا من المعالجة السريعة، والتي يصبح من الصعب أو المستحيل التعامل معها باستخدام قواعد البيانات التقليدية والأدوات التحليلية التقليدية، علما بأن معالجة بيانات بهذا الحجم يتطلب برمجيات تعمل بشكل متوازي على العشرات أو المئات، أو حتى الآلاف من الخوادم[7]. وهي تتدفق بسرعة كبيرة وبأشكال مختلفة قد تكون دقيقة وغير دقيقة، الا أنها تحتوي على قيمة كبيرة، وقيمة هذه البيانات تتمثل في تحويلها الى معلومات ومعرفة[8]

وهناك من يعرفها بأنها مجموعة كبيرة من البيانات التي تم تجميعها من مصادر مختلفة، هذه البيانات كبيرة ومعقدة بحيث يصعب معالجتها باستخدام أدوات إدارة قواعد بيانات عادية أو تطبيقات معالجة بيانات تقليدية، والتعقيد يأتي من كون البيانات ترد من مصادر مختلفة وبكمية كبيرة.

في حين يعرفها معهد ماكينزي العالمي بأنها: مجموعة البيانات التي تفوق حجم أو قدرة أدوات قواعد البيانات التقليدية من التقاط، وتخزين، وإدارة وتحليل تلك البيانات.

كما قامت جارتنر بتحديث تعريفها للبيانات الضخمة ليصبح كالتالي: البيانات الضخمة هي أصول معلومات كبيرة الحجم، عالية السرعة، و/أو عالية التنوع تتطلب أشكال جديدة من المعالجة لتعزيز عملية صنع القرار والفهم العميق وتحسين العملية.

2--تضخيم البيانات

منذ فجر التاريخ وحتى عام 2003 راكمت البشرية ما قدره 05 اكسابايت من البيانات والمعلومات التي تمثلت في الكتب والدواوين والوثائق، وكل ما تم تدوينه منذ عصر الانسان الحجري القديم ولكن ابتداء من سنة 2012 نمت هذه البيانات وتضاعفت بشكل مذهل ليصل معدل النمو الرقمي للبيانات الى 2.72 زيتابايت، وأصبحت الـ05 اكسابايت التي أنشأت منذ الخليقة تنشأ في يومين فقط، بينما سيتضاعف حجم البيانات الموجودة حاليا كل سنتين من الزمن.

بينما توقعت شركة IMC العالمية نحو العالم الرقمي بـ:10 أضعاف عام 2020 ليصل الى 44 تريليونGB، مشيرة الى أن حجم البيانات الرقمية في أجهزة الايباد لوحدها ستعطي ثلثي المسافة الى القمر أي 157.674 ميل ما يعادل 253.704 كم إذا تم مراكمة هذه البيانات فوق بعضها خلال سنة 2020.[9] والشكل رقم (02) يوضح زيادة حجم البيانات بين 2012-2017.

3-صفات البيانات الضخمة

البيانات الضخمة لها ثلاث صفات هي الحجم(volume)، التنوع(variety) والسرعة(velocity)، اختصارا يطلق على هذه الصفات (3s' v):

1-3 الحجم: يكون حجم البيانات كبير جدا ويقاس بوحدات القياس التي تم ذكرها في الجدول رقم"1"، وهذه البيانات تأتي من مصادر كثيرة (كلما زاد حجم البيانات كل ما كان هناك فائدة أكبر من تحليلها والعثور على معلومات مفيدة منها.

2-3 التنوع: تأتي البيانات من مصادر مختلفة وفي أشكال مختلفة، فهي تكون نصوص، أصوات، فيديوهات، صور، بيانات واردة من أجهزة رصد والكثير من المصادر.

3-3 السرعة: وهو معدل تدفق البيانات التي قد ترد في شكل حزم (batch)، وقد ترد أسبوعيا أو يوميا، أو كل ساعة، أو دقيقة وكل ثانية، أو كل ميلي ثانية.

يتم تمييز البيانات الضخمة من خلال الحجم، التنوع، والسرعة. ومن خلال دراسة الحجم الكبير للبيانات يمكن للشركات أن تفهم زبائنها بشكل أفضل، تخيل مثلاً البحث في بيانات مشتريات مليون شخص يتعامل مع متجر وول مارت، هذا البحث والتحليل في الكم الهائل من فواتير المشتريات وتكرار المشتريات وتنوعها، سيعطي معلومات مفيدة جداً للإدارة ومتخذي القرار.

وتختلف نوعية البيانات الكبيرة التي يمكن جمعها من البيئات المختلفة، فهناك بيانات كبيرة، وخبرات أكبر، وبيانات متعمقة متداخلة، بالإضافة الى التفاعلات الاجتماعية الكثيرة، وبيانات مفصلة عن أنشطة التعلم من نصوص ووسائط ومقاطع فيديو، وتختلف هذه البيانات في نوعيتها وعمقها.

المحور الثاني: الشركات التقنية

أولا- حول شركة جوجل

1- لمحة عن الشركة

كلمة Google هي تحريف لمصطلح رياضي يعني عدد كبير جدا عبارة عن رقم 1 متبوعا بمائة صفر، استخدمته شركة جوجل ليدل دلالة واضحة على مهمة شركة جوجل وهي أن تقوم بتنظيم كم هائل وعدد لانهائي من صفحات الانترنت المليئة بالمعلومات لكي تجعلها متاحة لكل مستخدم من مستخدمي الإنترنت الذين يبحثون عن المعلومات في شتى المجلات[10]. بمعنى أن رسالتها تتلخص في تنظيم معلومات العالم، وتمكين الناس من الوصول الى هذه المعلومات والاستفادة منها. والصورة رقم (01) توضح ذلك.

كانت الانطلاقة الأولى لجوجل عام 1996 من قبل سيرجي برين Sergey Brin ولاري بيج Larry Page، وهما لازال طالبين في جامعة ستانفورد Stanford University، وبدأت الفكرة من استرجاع بيانات متشابهة من بين كم كبير وهائل من البيانات.

2 شركة جوجل في أرقام

تعتبر شركة جوجل من أكبر الشركات العالمية التي تعمل في مجال الانترنت، تحتل المرتبة الأولى عالميا في مجال الدعاية على شبكات الانترنت، وهو أكثر محرك بحثي يتم استخدامه في العالم، تقريباً 92.12% من عمليات البحث على الإنترنت تتم عن طريق محرك بحث جوجل، أي 3.5 مليار عملية كل يوم، فيما يعني متوسط 40,000 عملية بحث كل ثانية. وهو أكثر موقع يزوره الناس في العالم، في العام الماضي(2021) وصل عدد الزيارات إلى 62.19 مليار زيارة لمختلف الأسباب.[11]

ثانيا- حول شركة ميتافيرس

1- لمحة عن الشركة

شركة ميتا فيرس (بالإنجليزية: .Metaverse)، تمارس النشاط التجاري باسم ميتا (Meta) والمعروفة سابقًا باسم شركة فيسبوك (.Facebook)، والذي يُعد شبكة اجتماعية استأثرت بقبول و تجاوب كبير من الناس خصوصا من الشباب في جميع أنحاء العالم، وهي لا تتعدى حدود مدونة شخصية في بداية نشأتها في شباط 2004، في جامعة (هارفارد) في الولايات المتحدة الأمريكية من قبل طالب متعثر في الدراسة يدعى " مارك زوكربرج" و كانت مدونته محصورة في بدايتها في نطاق الجامعة و بحدود أصدقاء زوكربيرج الطالب المهوس في برمجة الكمبيتور،و لم يخطر بباله هو و صديقين له أن هذه المدونة ستجتاح العالم الافتراضي بفترة زمنية قصيرة جدا.[12]

2- شركة ميتا في أرقام

قال مارك زوكربيرج مؤسس شركة ميتا: لقد كانت لدينا نهاية قوية لهذا العام، حيث استمر الأفراد والشركات في استخدام خدماتنا خلال هذه الأوقات الصعبة حيث بلغ معدل استخدام الميتا 2.8 مليار مستخدم شهريا في سنة 2021، حيث يقضي الفرد الواحد 19.5 ساعة على الموقع.[13]

ومن خلال الجدول رقم (02) نلاحظ أن الميتا يستخدم كوسيلة للتسويق والاعلان بشكل كبير والدليل هو النسب والأرقام الكبيرة الخاصة بكل منهما، فنجد مثلا 200 مليون شركة مصغرة تستخدم الميتا للتسويق، و800 مليون مستخدم شهري لـــ Marketplace و80 مليون صفحة متخصصة للتسويق، كما نلاحظ أن 44% من المستهلكين يعترفون بتأثرهم وتفاعلهم لما يتم الإعلان عنه وتسويقه أين نجد 26% من المستخدمين يشترون عبر موقع ميتا، الذي بلغت عائداته من الإعلانات 2020 مليار دولار.

المحور الثالث: إدارة البيانات الضخمة في شركتي جوجل وميتا

أولا- أهمية البيانات الضخمة بالنسبة للشركات التقنية

عبرت العديد من الشركات التقنية عن رضاها نتيجة استخدامها واستغلالها للبيانات الضخمة، وذلك وفقا لشركة اكسنترaccenture، وهي تنظر للبيانات الضخمة كمحفز لتحولها، إضافة إلى أن أكبر الشركات تزداد حجما أكثر مع

استغلالها للبيانات، على الرغم من أن البيانات الضخمة تشكل تحديا إلا أن الشركات التقنية الكبرى تعد من أكبر المستفيدين منها.

البيانات الضخمة تسمح للمؤسسات بتصميم منتجات وخدمات جديدة ومعينة لتلبية احتياجات معينة، وهذا النهج معروف في مجال التسويق وإدارة المخاطر، كما تتيح البيانات الضخمة بعمل منتجات جديدة وتقديم خدمات تتماشى ومتطلبات المستفيدين، وابتكار أعمال جديدة كليا، كما تستخدم الشركات المصنعة البيانات التي يتم الحصول عليها من مستخدمي منتجاتها من أجل تحسين وتطوير منتجاتها في المستقبل، وخلق عروض ما بعد البيع، فقد أصبح للشركات إمكانية هائلة لتوليد قيمة كبيرة من قواعد البيانات، كما أنه بإمكانها عن طريق تطبيق التقنيات المتقدمة لرفع الإنتاجية، وخفض التكاليف، وتقديم ابتكارات جديدة، وتحسين الكفاءة وجودة المنتجات، فقد أصبحت البيانات عاملا هاما من عوامل الإنتاج اليوم، كما أصبحت إمكانية جمع البيانات تتم بسهولة وخاصة مع كثرة استخدام الوسائط المتعددة، ووسائل الإعلام الاجتماعية، وانترنت الأشياء.

إن عملية تحليل البيانات الضخمة تسعى إلى تحسين عملية صنع القرار الإداري، ولتطبيق عملية تحليل البيانات الضخمة يتعين على الشركات تحديد المجموعة الوظيفية، وإيجاد أنماط حقيقية، وترجمة الرؤى التحليلية إلى معلومات تجارية فعالة للاستفادة القصوى من عملية تحليل البيانات الضخمة، بالإضافة إلى إعداد فريق من الموظفين والخبراء والمجهزين بالموارد والتقنيات اللازمة للتعامل مع مختلف الجوانب الخاصة بالكميات الكبيرة من البيانات وكذلك على دراية من أن تحليل البيانات الضخمة يحسن من العمليات الداخلية مثل: إدارة المخاطر، إدارة علاقات العملاء، والخدمات اللوجيستية، والاستفادة من المعلومات واستهداف العروض المناسبة لعملائها في الوقت المناسب.[14]

ثانيا-البيانات الضخمة للشركات التقنية

1-البيانات الضخمة في شركة جوجل

كانت جوجل هي بداية عصر الانتقال إلى البيانات الضخمة، التي تعالج عبر مجموعة من الأجهزة والخوادم. الان نعيش بداية حقبة جديدة تعرف بإنترنت الأشياء، شبكة ضخمة من الأشياء التي تمتلك عنوان IP بمقدورها إرسال وتبادل بيانات مع جميع الأجهزة المرتبطة بها.

أصبحت شركة جوجل بوابة الانترنت، حيث تتلقى مليار طلب بشكل يومي من 20 لغة مختلفة يتم تداولها عبر صفحة جوجل الرسمية.

تدير جوجل مجموعة من أكبر مراكز تخزين البيانات حول العالم، حيث أنها أطلقت موقعا خاصا بمراكز البيانات data centers لديها، ويقدم الموقع معلومات عن مراكز البيانات والية تبريدها والنسخ الاحتياطي، وأماكن تواجدها في العالم، والكثير يتساءل حول كم ومساحة مراكز تخزين هذه البيانات الضخمة التابعة لأكبر شركة لخدمات الانترنت في العالم وصاحبة لأكبر موقع الكتروني في العالم، وتوضح الصورة رقم (02) أهم مراكز تخزين البيانات الضخمة لجوجل.

تتميز مراكز تخزين البيانات الخاصة بشركة جوجل بنظام امان عال جدا، حيث يحتاج البصمة العين للدخول، وقبلها المرور بعدة مراكز تفتيش.

بدأت جوجل بحل جزء من مشكلة البيانات الضخمة في العام2003 وذلك بنشرها لورقة بعنوان نظام ملفات جوجل the google file system هذه الورقة قدمت حل لإمكانية كتابة وقراءة الملفات الموزعة بكفاءة عالية. ثم قدمت جوجل حل للجزء الأخر من المشكلة في عام 2004وذلك في الورقة التي نشرتها عن كيفة معالجة البيانات الضخمة في بيئة معالجة متوازية وعلى بيانات موزعة. وكانت الورقة بعنوان MapReduce: simplified data processing on large clusters.

تعالج جوجل عدد كبير جدا من البيانات يفوق ما لدى أية شركة على هذا الكوكب، وتميل جوجل لابتكار أدوات للتعامل مع هذه البيانات، وقد كشفت جوجل مؤخرا أنها تخلت عن نهج هادووب لأنه يميل للقيام بالعمليات على شكل دفعة، وهو مايعني أن يؤمر الكومبيوتر بتنفيذ عمليات بكميات كبيرة، ومن ثم انتظار النتيجة، وقد اقترحت جوجل خدمة السحاب لتدفق البيانات بالإضافة إلى عمل الهادوب تقوم أيضا بعمل وظائف تحليل البيانات على المعلومات الصحيحة أثناء وردها في قاعدة البيانات.

2-البيانات الضخمة في شركة ميتا

عندما نتحدث عن الميتا فإننا نتحدث عن عدد كبير جدا من المستخدمين ومقدار كبير جدا أيضا من البيانات الضخمة التي تستمر في التزايد في كل ساعة والتي بالطبع ستحتاج الى سيرفرات ومساحات تخزين كبيرة الحجم، لذلك قرر موقع الفيس بوك انشاء مركز بياناته الخاص الذي سيوفر له القدر

الكبير الذي يحتاجه لسد حاجاته من مساحة تخزين البيانات المتزايدة وسيوفر عليه الطاقة الكبيرة التي تستهلكها السيرفرات السابقة.

يمكن القول أن شركة ميتافيرس هي أكثر شبكات التواصل الاجتماعي شعبية على مستوى العالم، حيث يوجد بها أكثر من ملياري مستخدم نشط على مستوى العالم، حيث تخزن كميات هائلة من بيانات المستخدمين، مما يجعلها معجما عجيبا للبيانات، ويمثل الفيس بوك خامس أكبر شركة عامة في العالم، بقيمة سوقية تصل إلى 321 مليار دولار تقريبا.

تشير الإحصائيات إلى أن كل 60 ثانية يتم تحميل 136000 صورة، ويتم نشر 510.000 تعليقات... هذه المعلومات تعني الكثير، ولكن مع مثل هذه البيانات يعرف الفيس بوك من هم أصدقاؤنا، ومانبديه، وأين نحن، ومانفعله، ومانحبه...وأكثر من ذلك بكثير، ومن هنا فعلى الأرجح أن الميتا هو الشركة الوحيدة التي تمتلك هذا المستوى الرفيع من معلومات المستخدمين التفصيلية، فكلما زاد عدد المستخدمين زادت المعلومات التي يجمعونها.[15]

ثالثا- تقنيات البيانات الضخمة المستخدمة في شركتي جوجل و الميتا

1-هادووب

هادوب" هو برنامج أو منصة برمجية مفتوحة المصدر مكتوبة بلغة الجافا لتخزين ومعالجة البيانات الضخمة بشكل موزع مثل تخزين بيانات ضخمة على عدة أجهزة ومن ثم توزيع عملية المعالجة على هذه الأجهزة لتسريع نتيجة المعالجة. ويمكن الاستفادة من تحليل هذه الأنواع من البيانات الكبيرة في التعليم، لتوفير مجموعة متنوعة من الفرص والخيارات بهدف تحسين تعلم الطلاب وإضفاء الطابع الشخصي على مسار الطالب إلى إتقان المحتوى، من خلال التعلم التكيفي أو التعليم القائم على الكفاءة، مما ينتج عنه تعلم أفضل نتيجة لتشخيص أسرع وأكثر تعمقا لاحتياجات التعلم أو المتاعب التي تواجهه أثناء عملية التعلم، بما في ذلك تقييم المهارات مثل التفكير المنظم، والتعاون، وحل المشاكل في سياق عميق، و تقييم أصيل لمجال وموضوع المعرفة، بالإضافة لتحديد التدخلات المستهدفة لتحسين نجاح الطلاب وخفض التكاليف الإجمالية للطلاب والمؤسسات، واستخدام البيئات القائمة والمعلومات المعقدة في صنع القرارات وتحديد السياسات.

ويمكن أن توفر هذه البيانات أيضا أدوات حديثة وفعالة لقياس أداء الطلاب للمهام التعليمية، كما يُمكِّن قياس هذه الأنواع من المهام زيادة أهمية ودقة النتائج عن كيفية تعلم الطلاب، ويمكن أن تساعد كذلك في تصميم بيئات تعلم تصميما مخصصا وفق احتياجات محددة للطلاب، ويمكن أن تعطي تحليلا واضحا لردود الفعل الفردية والجماعية لمجموعة من القضايا التعليمية

يقول جاي بارخ نائب رئيس قسم هندسة البنية التحتية في الفيس بوك " يشغل الفيس بوك أكبر مجموعة في العالم من الهادوب، التي تتجاوز 4000 جهاز وتخزين أكثر من مئات الملايين من الجيجابايت، وتعمل هادوب بمعالجة السجلات ونظام التوصيات، وتخزين البيانات بالاعتماد على قاعدة البيانات هادوب مثال apache h base التي تحتوي على بنية متعددة الطبقات التي تدعم الرسائل في يوم واحد.

2- Scuba

وهي تقنية خاصة بالبيانات غير المهيكلة، وهي تساعد مطوري hadoop على الغوص في مجموعات البيانات الضخمة، يسمح scuba بتخزين البيانات المجمعة في الذاكرة، وتحليل فوري ومخصص لها، كما يقوم بجمع البيانات من مراكز البيانات المتعددة ووضعها وتنسيقها في بيانات السجل الموجود في نظام الذاكرة والتي يمكن الوصول إليها على الفور.

3- Apache hive

وهي بينية تحتية متكاملة لتخزين البيانات على hadoop، تمكن من التحليل والاستعلام عن طريق تركيب لغة SQL وتجميع البيانات من الناحية اللغوية.

4- Prism

شعر فريق الفيس بوك انه بحاجة الى العديد من مراكز البيانات، ومن هنا قاموا بتطوير prism، وهي عبارة عن منصة تقدم العديد من مساحات الأسماء بدلا من القاعدة الواحدة التي تحكمها hadoop.

خاتمة:

تعتبر إدارة البيانات الضخمة أحد أهم التطورات على المستوى الإداري، وواقعا جديدا أفرزته التطورات في عالم التكنولوجيا، حيث أصبح الاعتماد عليها مهمة الشركات التقنية نظرا لإمكانياتها الكبيرة، على غرار شركتي جوجل وميتا، وبعد الدراسة التحليلية لإدارة البيانات الضخمة في هذين الشركتين توصلنا للعديد من النتائج، أهمها :

- تقوم الشركات التقنية بمعالجة ملايين البيانات الرقمية، وتحويلها إلى معلومات تستفيد منها للتوصل إلى أخذ قرارات سليمة، إضافة إلى التنبؤ بالمستقبل.

- تستخدم الشركات التقنية بيانات المستخدمين في عمليات التسويق الخاصة بها، وذلك عن طريق جمع تلك البيانات وتحليلها واستخدام النتائج لأغراض تجارية.

- تمتلك شركتي جوجل وميتا بيانات أكثر من ثلث سكان الأرض.

- هناك العديد من التقنيات التي تستخدمها شركتي جوجل وميتا في معالجة البيانات، أشهرها هادووب.

- رأس مال هذه الشركات هو بيانات المستخدمين، لهذا توفر لهم خدمات مجانية متاحة للجميع، وذلك لتحصيل أكبر قدر من البيانات.

الجدول(1): وحدة قياس البيانات

وحدة القياس		الحجم	الوصف
بت	Bit	-	أصغر جزء يمكن قياسه ويتكون من رقم واحد
بايت	Byte	bit1024	كل(8) بت يكون حرف واحد
كيلوبايت	Kilobyte	1024 byte	يستخدم لقياس حجم الملفات
ميجابايت	Megabyte	1024 kilobyte	يستخدم لقياس أحجام اقراص التخزين ووحدات الذاكرة
جيجابايت	Gigabyte	1024 megabyte	يستخدم لقياس أحجام التخزين
تيرابايت	terabyte	1024 gegabyte	ذات السعة الكبيرة

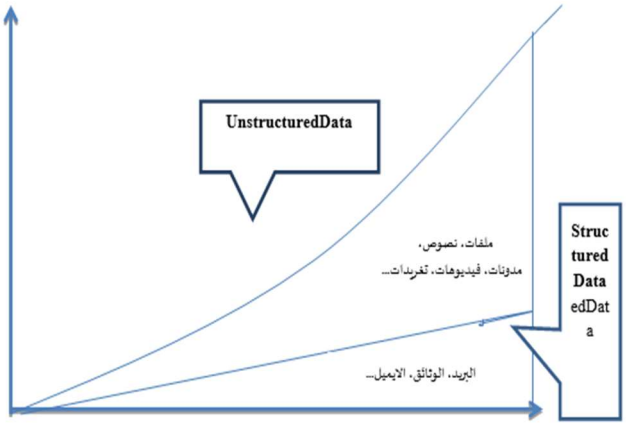
المصدر: http :www.computermt45,blogspot.com

الجدول(2): إحصائيات التسويق والإعلان على موقع ميتا (سنة 2021)

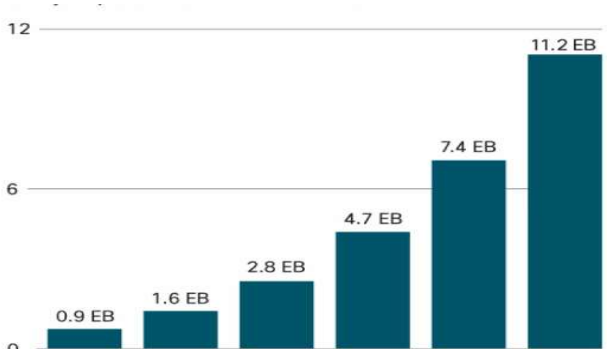
نوع الاستخدام	التسويق		الإعلان	
أرقام ونسب الاستخدام	الشركات المصغرة صفحة أعمال الاستخدام الشهري لMarketplace المستهلكين الذين يعترفون بتأثر سلوك التسوق	200 مليون 80 مليون 800 مليون %44	العائدات المشترون النقر على الإعلان	2020 مليار دولار %26 11 إعلانا شهريا

المصدر: Allgreen Matt, statistics and fact about: facebook for website hosting rating, http : //www. Websitehostingrating . com ,retrieved 27/03/2022,17.16

الشكل(1): البيانات المهيكلة وغير المهيكلة

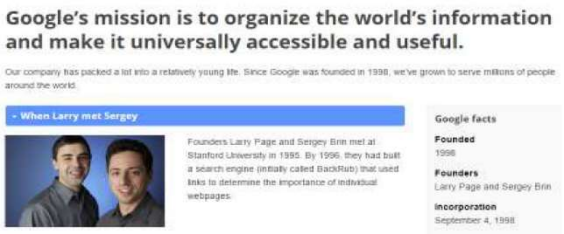


الشكل(2): زيادة حجم البيانات بين 2012-2017

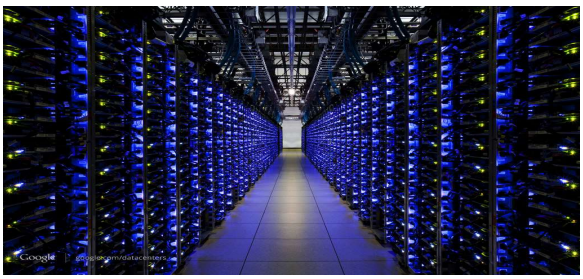


المصدر: تقرير مؤشر شبكة سيسكو المرئية عن حركة البيانات المتنقلة عالميا 2012-2017 متوفر على الخط: https://www.tech-wd.com

الصورة(1): رسالة جوجل حول نشاطها



الصورة(2): احدى مراكز بيانات جوجل



قائمة المراجع:

- **الكتب:**

- مؤيد سعيد السالم، تنظيم المنظمات: دراسة في تطوير الفكر مائة عام، دار الكتاب الحديث، عمان-الأردن، 2002.
- ايمان فاضل السامرائي، هيثم علي الزغبي، نظم المعلومات الإدارية، دار الصفاء، عمان-الأردن، 2004.
- عبد الفتاح المغربي، نظم إدارة المعلومات الإدارية، المكتبة العصرية، جامعة المنصورة مصر، 2002.
- زري تش، قصة فيس بوك: ثورة وثروة، ترجمة الهلالي وائل محمود محمد، القاهرة، 2003.
- كتاب ماذا تعرف عن جوجل، ص06.

- **الاطروحات:**

هناء قيراطي، أسامة دحمون، توظيف البيانات الضخمة في الشركات التقنية وخصوصية المستخدم، مذكرة لنيل شهادة الماستر في تكنولوجيا المعلومات والاتصالات، جامعة قالمة. 2017.

الملتقيات:

سعاد بوعناقة، البيانات الضخمة في قطاع المكتبات: نقاط القوة و الضعف والفرص والتهديدات، ورقة مقدمة ضمن فعاليات المؤتمر الرابع والعشرون لجمعية المكتبات المتخصصة حول: البيانات الضخمة وأفاق استثمارها: الطريق نحو التكامل المعرفي، مسقط، 6-7 مارس 2018،

- **مواقع الانترنيت:**

- بناي ضياء فالح، هل سمعت يوما بتحليل البيانات الكبيرة أو الضخمة، https://portal.arid.my/ar، 22/07/2022، 11.58.
- محمد نور، إحصائيات مذهلة حول البحث على جوجل واستخدامه https://www.alrab7on.com

- Allgreen Matt, statistics and fact about facebook for website hosting rating, http : //www. Websitehostingrating . com ,retrieved 27/03/2022,17.16
- Rolf sint, structured information in semantic wikis.
- Austria,http/ceur.ws.org/vol464/paper.p02

www.simplilearn.com: how face book using big data

1 بناي ضياء فالح، هل سمعت يوما بتحليل البيانات الكبيرة أو الضخمة، https://portal.arid.my/ar، 2022/07/22، 11.58.

2 مؤيد سعيد السالم، تنظيم المنظمات: دراسة في تطوير الفكر مائة عام، دار الكتاب الحديث، عمان-الأردن، ،2002، ص184.

3 ايمان فاضل السامرائي، هيثم علي الزغبي، نظم المعلومات الإدارية، ، دار الصفاء ، عمان-الأردن ،2004ص24.

4 عبد الفتاح المغربي، نظم إدارة المعلومات الإدارية، المكتبة العصرية، جامعة المنصورة، مصر،2002، ص29.

5 Rolf sint, structured information in semantic wikis, Austria,http/ceur.ws.org/vol464/paper.p02.

6 Sponsored report few,deep analytics, few, http/whitepapers/list whit epaper.

6 مشروع البيانات الضخمة (الإطار التنظيمي)، هيئة تنظيم الاتصالات، شبكة الهيئات العربية لتنظيم الاتصالات وتقنية المعلومات، سلطنة عمان،2015، ص06.

7 سعاد بوعناقة، البيانات الضخمة في قطاع المكتبات: نقاط القوة و الضعف والفرص والتهديدات، ورقة مقدمة ضمن فعاليات المؤتمر الرابع والعشرون لجمعية المكتبات المتخصصة حول: البيانات الضخمة وأفاق استثمارها: الطريق نحو التكامل المعرفي، مسقط، 6-7 مارس 2018، ص80.

8 ماذا يعني مصطلح big data ، متوفر على الخط: http/www.massarat.ma، يوم 2022/04/24 على الساعة 17.02.

10 وليد محمد، مالا تعرفه عن جوجل، 2009، ص06.

11 محمد نور، إحصائيات مذهلة حول البحث على جوجل واستخدامه https://www.alrab7on.com/

12.زري تش، قصة فيس بوك: ثورة و ثروة، ترجمة الهلالي وائل محمود محمد، القاهرة، 2003،ص132

[13] Allgreen Matt, statistics and fact about facebook for website hosting rating, http : //www. Websitehostingrating.com,retrieved 27/03/2022,17.16.

13 هناء قيراطي، أسامة دحمون، توظيف البيانات الضخمة في الشركات التقنية وخصوصية المستخدم، مذكرة لنيل شهادة الماستر في تكنولوجيا المعلومات والاتصالات، جامعة قالمة، 2017، ص50 .

14 how face book using big data www.simplilearn.com, 11/04/2022,15.12.