البيانات الضخمة: ماهيتها وأهميتها وعناصرها

د. الحسن شعبان أحمد محمد
وحدة المكتبة الرقمية – جامعة المنيا
elhassanshaban@gmail.com

المستخلص

تهدف هذه المقالة إلى التعرف على ماهية البيانات الضخمة وتحليلاتها وأهميتها وعناصرها، ورصد البيانات الضخمة بالمكتبات، وحصر التحديات التى تواجه تطبيق تحليلات البيانات الضخمة، والتعرف على أنواع تحليلات البيانات الضخمة، ومن أهم النتائج التى توصلت إليها الدراسة أن تحليل البيانات الضخمة يساعد على رصد واكتشاف أوجه القوة والضعف، ويقدم للمسئولين حلولًا للمشكلات وفقًا لنتائج التحليل، بما يتناسب مع طبيعة الجهة أو الهيئة، وأن تحليل البيانات الضخمة يزيد من فرص التنافس بما يحقق ميزات ونتائج جديدة، وتزيد من فرص التطور والتنمية وأن تحليل البيانات الضخمة يساعد في اتخاذ قرارات صحيحة، وزيادة فرص التنبؤ بالمستقبل، وتتمثل أدوات تحليل البيانات الضخمة فى ثلاث أدوات رئيسة وهي أدوات التنقيب وأدوات التحليل وأدوات عرض النتائج، وأن بيانات المكتبة تعتبر بيانات ضخمة وذلك بسبب مقتنياتها ذات التنوع الواضح، والسرعة العالية، والحجم الهائل، ويمكن الاستفادة من تحليلات البيانات الضخمة فى المكتبة فى إدارة مقتنيات المكتبة وقواعد بيانات المستفيدين وتنمية الموارد البشرية بالاضافة لخدمات المعلومات المتاحة بالمكتبة وفى ترشيد سياسة الانفاق.

الكلمات المفتاحية: البيانات الضخمة – البيانات المهيكلة – أخصائى المكتبات والمعلومات – التحليل الرباعى.

تاريخ استلام البحث: 2022/1/22 تاريخ إجازة البحث: 2022/2/4

0/ تمهيد

لا شك أن البيانات والمعلومات المتاحة تتسع حاليًا بشكل متسارع جدًا، وتتنوع مصادرها تنوعًا كبيرًا، بالإضافة إلى أنها تُتاح فى أشكال وأنماط مختلفة، وللبيانات المنظمة والدقيقة دور كبير فى اتخاذ القرارات السليمة، ورسم السياسات الاستراتيجية، لدى متخذي القرار، وبالرغم من أن وفرة البيانات لمن يطلبها بشكل سريع يعتبر ميزة هذا العصر، إلا أنه فى الوقت نفسه يضع تحديات ضخمة لمتخذى القرار، حيث تعتبر البيانات الضخمة أحد أكبر التحديات التى تواجهها الشركات والمؤسسات الكبيرة، بسبب صعوبة معالجة هذه البيانات، والاستفادة منها (السالمى و عرابه، 2018).

وتعتبر البيانات أحد أهم الأدوات الرئيسة لأداء أنشطة المؤسسات اليومية، وعلى رأسها المكتبات ومراكز المعلومات، ونظراً لأهميتها فقد خُصص لها علم قائم بذاته هو "علم البيانات" "Data science" ذلك العلم هو المَعنىّ باستخراج أكبر قدر من المعرفة بأقل بيانات متاحة.

ويشير "كولتاى" إلى ضرورة العمل على محو الأمية فى هذا الإطار الذى صار الموضوع الأكثر تداولًا فى الأُوساط العلمية فى ضوء تبادل البيانات، وضبط جودتها، وإدارتها وتنظيمها، والاستشهاد بها، حيث تتصل معظم هذه الأنشطة السالف ذكرها بأنشطة المكتبات وأخصائى المعلومات. (Koltay, 2012)

1/ مشكلة الدراسة

يعيش قطاع المعلومات عامة وقطاع المكتبات خاصة على وقع نمو غير مسبوق للبيانات، نتج عنه ما يعرف حالياً بظاهرة البيانات الضخمة ، نتيجة للزيادة الهائلة في عدد الأجهزة المتنقلة وانتشار تطبيقات الحوسبة السحابية والمشاركة على منصات التواصل الاجتماعي، وتوفر هذه البيانات فرص واسعة للمكتبات لتطوير قيمتها التنافسية ورفع كفاءة خدماتها، وهذا يستلزم التعريف بالبيانات الضخمة وأهميتها وعناصرها وتحليلاتها وأنواعها.

2/ أهمية الدراسة

تستمد هذه الدراسة أهميتها من أهمية تحليلات البيانات الضخمة وما تقدمه من حلول ونتائج تساهم فى تطوير الخدمات وترشيد سياسة الإنفاق بما يتلاءم مع طبيعة الاستخدام.

3/ أهداف الدراسة

تهدف الدراسة إلى :

1. التعرف على ماهية البيانات الضخمة وتحليلاتها.
2. التعرف على أهمية تحليلات البيانات الضخمة وعناصر تحليلات البيانات الضخمة وأدوات التعامل مع البيانات الضخمة.
3. التعرف على أهمية البيانات الضخمة وتحليلاتها واستخداماتها فى مجال المكتبات والمعلومات.

4/ تساؤلات الدراسة

تُحاول الدراسة الإجابة على التساؤلات الآتية:

1. ماهية البيانات الضخمة وتحليلاتها؟
2. ما أهمية البيانات الضخمة وتحليلاتها؟
3. ما عناصر وأدوات تحليلات البيانات الضخمة؟

5/ مصطلحات الدراسة

1/5 البيانات الضخمة Big Data

مجموعة من البيانات التى لايمكن الحصول عليها وتخزينها وإدارتها بواسطة برمجيات قواعد البيانات التقليدية وأن حجم هذه البيانات تنمو بمرور الوقت أو بالتقدم التكنولوجى وتظهر فى تطبيقات مختلفة عن بعضها البعض.(Chen & Liu, 2014)

التعريف الإجرائى

البيانات الضخمة هى بيانات ذات حجم كبير ولا يمكن معالجتها بالطرق التقليدية وتتطلب قدرًا من المعالجة السريعة وتقنيات حديثة لمعالجتها وتحليلها.

2/5 تحليلات البيانات الضخمة Big Data Analytics

هى عملية معقدة لفحص مجموعات البيانات الكبيرة والمتنوعة أو البيانات الضخمة للكشف عن المعلومات بما في ذلك الأنماط المخفية والارتباطات غير المعروفة واتجاهات السوق وتفضيلات العملاء التى يمكن أن تساعد على اتخاذ قرارات عمل مستنيرة. (Rouse, 2018)

3/5 علم البيانات Data Science :

هو علم دراسة البيانات ويتضمن تطوير أساليب لتخزين وتحليل البيانات لاستخراج معلومات مفيدة، ويهدف علم البيانات إلى اكتساب رؤى ومعارف من البيانات بكلا نوعيها المهيكلة وغير المهيكلة. (The Tech Terms Computer Dictionary)

4/5 البيانات المهيكلة Structured Data

هى تلك البيانات التى تسكن نماذج البيانات وهى بذلك خضعت لتنظيم ما وفقًا لمعرفات أو محددات منطقية بحيث تسمح بتحديد أجزاء معينة من المعلومات مما يسهل عملية التعرف عليها والبحث عنها واسترجاعها كما فى قواعد البيانات العلائقية. (النشرتى، 2014)

5/5 البيانات غير المهيكلة Un Structured Data

هى البيانات التى لم تخضع لأية صيغة من الهيكلة أو التنظيم لها بشكل مسبق عن إنشائها كما لا يظهر فيها المعرفات أو المؤشرات مما يجعل طبيعتها بمثابة كتلة من البيانات المجمعة لها بداية ونهاية دون أية إشارة إلى أقسامها أو القطاعات التى تشتمل عليها وتمثل هذه الفئة 85% من إجمالى محتوى الويب. (النشرتى، 2014)

6/5 الحوسبة السحابية Cloud Computing

هى تكنولوجيا تستخدم كخدمة تسمح بتوفير البرامج للوصول إلى البيانات وتخزينها لأجهزة الكمبيوتر عبر الشبكة كخدمة تكنولوجيا المعلومات. (Reitz, 2002)

7/5 التحليل الرباعى SWOT

هو تحليل للبيئة التى سيتم فيها وضع الخطة الاستراتيجية وتطبيقها وتعمد إلى تكوين قاعدة معلوماتية يستند عليها فى التخطيط الاستراتيجى. (أبوالنصر، 2012)

6/ الدراسات السابقة

وشرع الباحث بمسح الإنتاج الفكري العربي والأجنبي (التقليدي / الإلكتروني) والبحث في محركات البحث المختلفة وقواعد البيانات وفهارس المكتبات وفيما يلي أهم المصادر بحث الإنتاج الفكري :

- البحث في محرك البحث Google Scholar .
- البحث في فهارس مكتبات الجامعات المصرية من خلال الفهرس الموحد للجامعات المصرية (اتحاد مكتبات الجامعات المصرية)
- البحث من خلال المستودعات الرقمية المتخصصة بمجال المكتبات والمعلومات ومنها E-LIS، DLIST ، Open-doar

(Holt,2005 ; *The University of Arizona Libraries Repository;* Open DOAR)

- البحث في بنك المعرفة المصري www.ekb.eg

<u>ومن أهم الدراسات السابقة التى تم التوصل اليها ما يلى:</u>

دراسة (القصاص، 2019) هدفت هذه الدراسة إلى التعرف على ماهية البيانات الضخمة ونشأتها وتطورها واهميتها، ودراسة الاساليب والادوات والتقنيات المستخدمة فى تجميع البيانات الضخمة وتحليلها، وتحليل مكونات مشروعات مختلفة للبيانات الضخمة فى المكتبات، واستنباط مجالات جديدة فى سوق العمل تعتمد على البيانات الضخمة وتقنياتها، واستكشاف مدى ارتباط البيانات الضخمة بمقررات علوم المكتبات التى يتم تدريسها لأجيال أخصائى

المكتبات، واعتمدت الدراسة على المنهج الاستكشافى وذلك بالاعتماد على أسلوب البحث الوثائقى وأسلوب التحليل، ومن أهم النتائج التى توصلت إليها الدراسة أن البيانات الضخمة فرضت نفسها كواقع جديد دون تخطيط وذلك نتيجة للتطورات التكنولوجية السريعة، يمكن استخدام البيانات الضخمة لتكوين مؤشرات كلية تستخدم من قبل الدولة فى تحديد أهداف التنمية وتحديد أولويات استهداف المناطق المختلفة، وأن إطار البيانات الضخمة يحتاج إلى تعديلات تشريعية تخص الآليات التى يتم بها مشاركة البيانات، وأن من أفضل التقنيات التى تتعامل مع البيانات الضخمة ال Hadoop ، ومن أهم توصيات الدراسة ضرورة تغيير أسلوب إدارة البيانات بالمكتبات وذلك من خلال تحليلات البيانات الضخمة لسرعة اتخاذ القرار والتنبؤ لوضع المكتبة الحالى والمستقبلى، إعادة توظيف الهيكل التنظيمى بالمكتبات بحيث يتم إضافة قسم جديد خاص بدعم اتخاذ القرارات الحالية والمستقبلية، وعقد دورات تدريبية لأخصائى المكتبات لإكسابهم مهارات التعامل مع البيانات الضخمة وجمعها وتحليلها، والاستفادة من شبكات التواصل الاجتماعى كأداة لجمع البيانات الضخمة من أجل تحليلها للاستفادة منها في تحسين واقع المكتبات.

دراسة (اليحيائية، 2019) وهدفت الدراسة إلى التعرف عن قرب عن البيانات الضخمة وخصائصها أو سماتها، بالاضافة إلى التعرف على التحديات التى تواجه البيانات الضخمة فى ظل التطور الحاصل فى تكنولوجيا المعلومات والانفجار الرقمى، واعتمدت الدراسة على المنهج الوصفى، ومن أهم النتائج التى توصلت إليها الدراسة أن البيانات الضخمة علم حديث ومتجدد ويحتاج إلى موارد بشرية تستفيد منه ولكن قبل ذلك لابد من إيجاد المؤسسات التعليمية المتخصصة فى مثل هذه العلوم فالبيانات الضخمة تحتاج إلى مهارات فى تكنولوجيا المعلومات والاتصالات والتحليل وغيرها من العلوم الأخرى، وحققت الشركات نجاحًا خلال استغلال تحليلات البيانات الضخمة فى تسويق منتجاتها وتوسيع نطاق مشترياتها، ومن أهم توصيات الدراسة التطرق أكثر إلى البيانات الضخمة فى الدول العربية والتوجه نحو هذا الموضوع فى الأبحاث والدراسات لتذكير المؤسسات والشركات بأهمية البيانات الضخمة، على المؤسسات التعليمية وضع خطط واستراتيجيات نحو إدخال هذا العلم ضمن تخصصها والتركيز عليه.

دراسة (الأكلبى، 2018) هدفت هذه الدراسة إلى التعرف على أهمية تحليل البيانات الضخمة، والتعرف على كيفية الإفادة من البيانات الضخمة فى جامعة الملك سعود فى اتخاذ القرار، واقتراح نماذج تنبؤية ومؤشرات إضافية كأدوات للمساعدة فى إدارة البيانات الضخمة والسيطرة عليها بواسطة نظام "اتقان"، واعتمدت الدراسة على المنهج الوصفى التحليلى، ومن أهم النتائج التى توصلت إليها الدراسة أن المنظمات تمتلك الكثير من البيانات وبكميات ضخمة وغير منظمة فى الاغلب وهو مايجعلها تفقد الكثير من عوائد هذه البيانات لأنها غير مصنفة وغير منظمة ولا تمتلك هذه الجهات برمجيات متخصصة فى إدارة محتوى تلك البيانات الضخمة وتحليله، وأن متخذ القرار يحتاج بانتظام إلى توفر نتائج تحليل البيانات الضخمة لتكون ملائمه له وتناسب حاجاته عند اتخاذه للقرار، ولا زالت فائدة جامعة الملك سعود من خدمات نظام "اتقان" فى تحليل كافة البيانات الضخمة للجامعة دون المأمول مقارنة بقدرات النظام، وأن البيانات الضخمة تزداد ضخامة بسرعة هائلة وتحتاج إلى خطة معالجة استراتيجية على مستوى الجامعة، وأن من أهم أنظمة تحليل البيانات الضخمة نظام Hadoop وهو نظام مفتوح المصدر، وأن تقنيات الحوسبة السحابية من أهم الحلول والطرق المفيدة فى توفير مساحات تخزينية للبيانات الضخمة والتعامل معها بسرعة ، ومن أهم توصيات الدراسة ضرورة سن الأنظمة والتشريعات المنظمة على المستوى الوطنى والإقليمى والدولى لموضوع الملكية الفكرية وخصوصية البيانات، وإجراء دراسات حول انتهاك خصوصية البيانات وأضرارها، وإعداد خطة استراتيجية لاستيعاب الزيادة المستمرة والمتسارعة للبيانات الضخمة والعمل على تحليلها والاستفادة منها.

دراسة (على، 2018) هدفت الدراسة إلى التعرف على البيانات الضخمة ونشأتها ومصادرها وأنواعها، وتحديد دور أخصائى المعلومات والمؤهلات الواجب توافرها لديه، وتحليل أنواع البيانات الصحية الضخمة ومصادرها وخصائصها، وبيان أهمية وخطوات التحول بنظم الرعاية الصحية من الإدارة التقليدية إلى الإدارة بتحليلات البيانات الضخمة، وإلقاء الضوء على أهم التجارب والمشروعات الدولية التى نجحت فى إدارة البيانات الصحية الضخمة، وتقييم جدوى تطبيق تقنيات البيانات الضخمة فى مؤسسات الرعاية الصحية المصرية، واعتمدت الدراسة على منهجين وهما المنهج الوصفى التحليلى وأسلوب البحث الميدانى، ومن أهم النتائج التى توصلت إليها الدراسة أن البيانات الضخمة أثرت على مؤسسات الرعاية الصحية من حيث

مهارات العاملين وظهور وظائف جديدة لأخصائى المعلومات، وتتعدد مصادر البيانات الضخمة الصحية فى قطاع الرعاية الصحية مابين السجلات الصحية الإلكترونية والبيانات الإدارية، لا يوجد فى مستشفيات القطاع الحكومى وشبة الحكومى أى استخدام لتحليلات البيانات الضخمة لأى غرض كان، ومن أهم توصيات الدراسة عمل دبلومات متخصصة فى تحليل البيانات بكليات الحاسبات والمعلومات وأقسام المكتبات والمعلومات لإتاحة تحويل مسار العمل لأخصائى المعلومات أو الأطباء الراغبين فى التعمق والتخصص فى علوم البيانات، وتقديم منح تنافسية من المؤسسات الأكاديمية فى تدريب أعضاء هيئة التدريس والباحثين على أحدث تقنيات علوم البيانات، والتعجيل بتبنى تقنيات تعتمد على علوم البيانات والبيانات الضخمة.

تقريروزارة الاتصالات القطرية (2014) وهدف هذا التقرير إلى تعزيز وعى القطاع الخاص والعام بالتحديات والفرص التى تتيحها البيانات الضخمة بالاضافة إلى التأثير على العمليات التى تتعلق بوضع السياسات وتوفير الخدمات فى دولة قطر وما يتعلق بالمخاطر والتحديات فى قضايا الخصوصية وسرية تحليل البيانات، ومن أهم نتائج التقرير ضرورة التأكد بما لا يدع مجالاً للشك من قبل المؤسسات إلى أن تطبيق البيانات الضخمة لن يؤثر على خصوصية وسلامة سكان قطر، وضرورة تطبيق قانون حماية البيانات والخصوصية واستخدام الحلول التقنية المناسبة وعلى المؤسسات التعليمية والمؤسسات ذات الصلة أن تبذل مجهودا فى توفير مواد تعليمية تدريبية ملائمة لإدارة هذا العلم الجديد.

دراسة (Al-Barashdi, 2019) هدفت الدراسة إلى استكشاف التقنيات والأدوات المستخدمة فى تحليل البيانات الضخمة فى المكتبات الاكاديمية والتعرف على مستقبل التطورات البحثية فى البيانات الضخمة، والتعرف على ثغرات الدراسات البحثية حول البيانات الضخمة بالمكتبات الأكاديمية، وتوفير فهم أكثر لمزايا البيانات الضخمة بالمكتبات الاكاديمية، واعتمدت الدراسة على المنهج المسحى، ومن أهم النتائج التى توصلت إليها الدراسة تقنية البيانات الضخمة تسهل التعامل مع مجموعة البيانات وربطها للتنبؤ بالمستقبل وتقييم المخاطر، حيث يلعب أخصائى المكتبات دورًا مهمًا فى منظومة تحليل البيانات الضخمة من حيث تلبية احتياجات المستفيدين وتعزيز الاتصال بين المستفيدين والمكتبة، ومن أهم توصيات

الدراسة تسليط الضوء على التحديات التى تواجه تطبيق تحليلات البيانات الضخمة بالمكتبات الأكاديمية، والتركيز على التقنيات المستخدمة فى تحليل البيانات الضخمة.

دراسة (Kamupunga, 2019) هدفت هذه الدراسة إلى التعرف على أدوات تحليل البيانات الضخمة والتقنيات المناسبة للمكتبات الأكاديمية، والتعرف على البيانات الضخمة بالمكتبات الأكاديمية، والتعرف على فوائد تطبيق تحليلات البيانات الضخمة بالمكتبات الأكاديمية، ورصد الثغرات والعقبات الموجودة بالمكتبات الأكاديمية والتى تعيق تطبيق البيانات الضخمة، ومن أهم النتائج التى توصلت اليها الدراسة أن المكتبات الأكاديمية التى تستخدم تحليلات البيانات الضخمة هى أكثر إنتاجية وكفاءة من غيرها، وأن أدوات تحليل البيانات الضخمة تتطلب مهارات خاصة لتسهيل تحسين الأداء فى العالم التنافسى، ومن أهم توصيات الدراسة توجيه الطلاب والباحثن للاهتمام بالبيانات الضخمة وإجراء دراسات لتحديد كيف يمكن للبيانات الضخمة التأثير على القيمة الاقتصادية فى المكتبات الأكاديمية.

دراسة (Jharotia, 2019) هدفت الدراسة إلى التعرف على ماهية البيانات الضخمة وخصائصها، والهدف من تطبيق تحليلات البيانات الضخمة، ودور أمناء المكتبات فى ظل البيانات الضخمة، ومن أهم النتائج التى توصلت إليها الدراسة أن استخدام البيانات الضخمة مفيد ويسمح بكفاءة العمل بالاضافة لتقليل التكلفة ودعم الابتكار، وأن معالجة البيانات لابد أن يتم بأدوات متقدمة لاكتشاف المعلومات القيمة، وأن البيانات الضخمة تؤدى تصورات معقدة فيما يتعلق بالأنماط والاتجاهات فى السلوك البشرى، وأن البيانات الضخمة تتوسع باستمرار فى ثلاثة جهات الحجم والسرعة والتنوع، وأن أول من تبنى تحليل البيانات الضخمة هى الشركات ثم يليه قطاع التعليم، وتحتل المكتبات مركز الصدارة فى مبادرات دعم التنقيب عن البيانات ومعالجتها.

دراسة (Sonawane, 2018) ناقشت هذه المقالة تعريف البيانات الضخمة، وخصائصها، ومهارات أخصائى المكتبة وأدوات تطبيق البيانات الضخمة بالمكتبات، ومن أهم النتائج التى توصلت إليها الدراسة أن البيانات الضخمة استحوذت على انتباه علماء وتخصص المكتبات، وأن أخصائي المكتبات والمعلومات يمكن أن يلعبوا دورًا كبيرًا فى عالم البيانات الضخمة، ويمكن

لأخصائى المكتبات الاطلاع على التحليلات التى توفرها تقنيات البيانات الضخمة مما سيضيف قيمة للخدمات المختلفة التى يقدمونها.

دراسة (Li et al., 2017) هدفت الدراسة إلى تقديم نظرة شاملة للبيانات الضخمة وإطار تطبيقها فى بيئة المكتبات، ومناقشة مفهوم البيانات الضخمة ورصد التحديات والفرص المتاحة لتطبيق البيانات الضخمة بالمكتبات، واعتمدت الدراسة على المنهج الاستقرائى، ومن أهم النتائج التى توصلت إليها الدراسة أن معظم المكتبات تؤكد على جدوى تطبيق البيانات الضخمة وتميل إلى تطبيقها فى المكتبات، وأن العقبات الرئيسة مقابل تطبيق البيانات الضخمة بالمكتبات تكمن فى الموارد البشرية ومستوى تكنولوجيا المعلومات، ومن أهم توصيات الدراسة ضرورة توسيع فرص تطبيق البيانات الضخمة بالمكتبات وتعزيز مهارات أخصائى المكتبات، وتعزيز خدمات الإعارة بين المكتبات.

دراسة (Mutula, 2016) وهدفت هذه الدراسة إلى التعرف على تأثير صناعة البيانات الضخمة على علوم المكتبات والمعلومات بما فى ذلك مناقشة فوائد استخدام تحليلات البيانات الضخمة فى جميع المجالات الأكاديمية وتزايد النمو والاهتمام بالبيانات الضخمة بسبب انخفاض تكلفة تخزين البيانات، ومن أهم النتائج التى توصلت إليها الدراسة أن البيانات الضخمة لها تأثير هائل فى جميع المجالات الأكاديمية والتى ساعدت فى خلق مهارات جديدة فى مختلف القطاعات الأكاديمية بما فى ذلك علوم المعلومات فى مجال إدارة البيانات والتنظيم والأرشفة والبحث والاسترجاع.

دراسة (Sengupta, 2016) هدفت هذه الدراسة إلى التعرف على ماهية البيانات الضخمة، واستعراض خصائص البيانات الضخمة، والتعرف على المنصات التى تستخدم تحليلات البيانات الضخمة، ومن أهم النتائج التى توصلت إليها الدراسة أن استخدام واعتماد البيانات الضخمة يتميز بالكفاءة بالاضافة لدقة النتائج والسماح بالابتكار، وأن أخصائى المكتبات لا يحتاج إلى أن يكون مبرمجًا ولكن بكيفية معرفة كيف يمكن بأدوات البرامج المختلفة تحويل كميات هائلة من البيانات إلى بيانات قيمة ومفيدة، البيانات الضخمة بالمكتبات أتاحت الفرصة لأخصائى المكتبات للعب دور مهم فى عالم البيانات الضخمة لأن لديهم المهارات والمعرفة والقدرة على مساعدة الجميع.

دراسة (stryk, 2015) تناولت هذه الدراسة موضوع البيانات الضخمة من جهة التعريف بأفضل ممارسات المسؤولين عن معالجة البيانات الضخمة التى تم الاطلاع عليها وآليات تحليلها واستخراج المنافع منها، واعتمدت الدراسة على أسلوب دلفى وهو أسلوب تفاعلى يعتمد على التواصل مع مجموعة من الخبراء من أجل استطلاع آرائهم حول موضوع معين، والتعرف على كيفية قيام المنظمات بإعداد البيانات وتنظيمها ليتم استخدامها بشكل مفيد فى إدارة المنظمة من خلال مساعدة المنظمات على استخراج معلومات أكثر فعالية للحصول على قيمة مضافة لها، ومن أهم نتائج الدراسة أن السبب الرئيس الذى دفع المنظمات للقيام بمعالجة البيانات الضخمة وتحليلها والاستثمار فى ذلك لأجل تقليل المخاطر التنظيمية والتقنية وتحقيق أرباح أعلى من هذه البيانات، وأن تنظيم البيانات وتحليلها يساعد فى تقديم معلومات مفيدة تساند المسؤولين فى اتخاذ القرار وضمان مستقبل قوى للمنظمة.

دراسة (Fister, 2015) وناقشت هذه الدراسة المسائل المتعلقة بتأثير البيانات الضخمة على مسألة الخصوصية التى تتعلق بحياة الناس، ودور المكتبات الأكاديمية فى مواجهة مسألة الخصوصية حيث تواجه هذه المكتبات العديد من التحديات فى عصر البيانات الضخمة سوف يترتب عليها دعم استخدام البيانات والمحافظة عليها باعتبارها ذات قيمة متزايدة فى النظام البيئى للمعرفة، بالإضافة إلى الاستفادة من تحليل البيانات الضخمة فى تنمية مجموعات المكتبة وخدماتها.

دراسة (Zakir, 2015) هدفت هذه الدراسة فى المقام الأول إلى التعرف على التقنيات التى تستخدم فى تحليلات البيانات الضخمة والتى تعمل معًا للمساعدة فى التوقعات المستقبلية، وتحليل استخدام تحليلات البيانات الضخمة وتأثيرها على القيمة التجارية للمؤسسات والفرص التنافسية، ومن أهم النتائج التى توصلت إليها الدراسة أنه على الرغم من أن تقنيات تحليلات البيانات الضخمة لاتزال فى مراحلها الأولى إلا أنه يتم الضغط على المنظمات لتطبيقها، وأن تحليلات البيانات الضخمة تمكن من الكشف عن أنماط واتجاهات جديدة واكتساب رؤى إضافية ، وأن تحليلات البيانات الضخمة تساهم فى خفض التكاليف وتحسين الإنتاج والكشف عن فرص جديدة، بالإضافة إلى اتخاذ قرارات أفضل اعتمادًا على البيانات.

دراسة (chen&zhang, 2014) هدفت هذه الدراسة إلى التعريف بالبيانات الضخمة وأهم التطبيقات المستخدمة فى تخزينها، والتعرف على التحديات والتقنيات المستخدمة فى إدارة البيانات الضخمة، ومن أهم النتائج التى توصلت إليها الدراسة أن مسألة البيانات الضخمة صارت قضية مؤرقة وأضحت محل اهتمام الكثير من الباحثين وصناع القرار فى القطاعات الحكومية وفى الشركات بشكل أكبر، وأن البيانات الضخمة سببت متاعب كبيرة لصعوبة الاستفادة منها وعدم القدرة على التحكم فيها، وأن البيانات الضخمة تحوى على الكثير من الكنوز والمنافع غير المستغلة والتى لم يتم الكشف عنها عبر تحليلها واستخلاص النتائج منها.

دراسة (Zicari, 2013) تناولت هذه الدراسة تحديات البيانات الضخمة وآفاقها المستقبلية وركزت بشكل خاص على الخصائص الرئيسة للبيانات الضخمة والتحديات المرتبطة بها، ومن أهم نتائج الدراسة أنها صنفت التحديات إلى ثلاثة أقسام : الأول: التحديات المرتبطة بالبيانات وأهمها الحجم والنوع والسرعة والمصداقية وجودة البيانات وتوافرها واكتشاف البيانات وملائمتها وشموليتها والخصوصية والهيمنة؛ والثانى : تحديات المعالجة وتشمل أسر البيانات وصعوبة الاختيار من بين البيانات المتشابهة وتحويل البيانات إلى شكل قابل للتحليل ونمذجة البيانات وفهم المخرجات وتصويرها وتقاسم النتائج وكيفية عرض التحليلات المعقدة للبيانات ؛ أما القسم الثالث: من التحديات فيتعلق بالإدارة وتشمل خصوصية البيانات وأمنها وحوكمتها وأخلاقياتها.

دراسة (Ularu. 2012) هدفت الدراسة إلى التعرف على ماهية البيانات الضخمة، وتحليلات البيانات الضخمة، وأهمية تحليلات البيانات الضخمة، وتحديات البيانات الضخمة، وبرامج تحليلات البيانات الضخمة، ومن أهم النتائج التى توصلت إليها الدراسة أن الشركات هى أول من استخدمت تحليلات البيانات الضخمة وذلك من أجل فهم وجذب العملاء وزيادة المبيعات وفهم العملاء، وأن من أهم وأشهر أدوات تحليلات البيانات الضخمة الهادوب، وأن من أهم فوائد تحليلات البيانات الضخمة أنها تساعد على اتخاذ قرارات صحيحة مبينية على البيانات.

7/ نبذة تاريخية عن مصطلح البيانات الضخمة وتطوره

ظهر مصطلح البيانات الضخمة فى بدايات عام 2000م، وزادت استخداماته –حديثًا- حيث انتشر المصطلح ضمن أوائل الاتجاهات التقنية المقبلة، إذ توقعت أهميتها وانتشارها - مراكز الأبحاث التقنية مثل : جارتنر، وماكينزى، وشركة IBM. ولم يخلُ الأمر من اهتمام دوائر سياسية كبيرة بهذا الموضوع التقنى، مثل: إدارة الرئيس الأمريكى السابق، والمفوضية الأوروبية التى عدت البيانات الضخمة رصيدًا أساسيًا للاقتصاد، كما ركزت دوائر علمية أبحاثها فى هذا المجال مثل: مؤسسة العلوم الوطنية الأمريكية، ومجلس الأبحاث الهندسية العلوم الطبيعية الكندى، ومعهد مهندسى الكهرباء والإلكترونيات الأمريكى، وبرنامج الأبحاث والابتكار الأوروبى، ومجلة الطبيعة، ومجلة العلوم، وقطاع الأعمال والاقتصاد مثل: منتدى "دافوس الاقتصادى"، كما انتشر الحديث عن البيانات الضخمة فى وسائل الإعلام مثل: نيويورك تايمز، وول ستريت جورنال، والايكونومست، ومجلة الفيصل العلمية.

فى عام 2007م كانت البشرية قادرة على تخزين (290) إكسا بايت من البيانات المضغوطة، وتنفيذ (6.4) إكسا بايت من التعليمات فى الثانية على أجهزة الحاسب، ولكن حتى الآن لم يتم اكتشاف سوى جزء صغير جدًا من قيمة البيانات التى تم إنتاجها من خلال استخدام تحليلات البيانات الضخمة. (الرباعى والطيب، 2018)

تتعدد مصادر المعلومات الرقمية الحالية سواء أكانت مقاطع فيديو أو الصور المتواجدة بالهواتف الذكية والتى يتم تحميلها على اليوتيوب، بالإضافة إلى البيانات المصرفية، وأجهزة الصراف الآلى، وتسجيلات كاميرات الأمن فى المطارات والفاعليات الكبرى، فضلًا عن بيانات الاصطدام الذرى التى سجلتها منظمة الأبحاث النووية الأوربية (سيرن).

وباستخدام تحليلات البيانات الضخمة يمكن أن نكتشف أنماط الاستخدام فى وسائل الإعلام الاجتماعية، والعلاقات المتبادلة فى البيانات العلمية والطبية من الدراسات المنفصلة وتقاطع المعلومات الطبية مع البيانات الاجتماعية.

8/ ماهية البيانات الضخمة

مع الزيادة المستمرة فى كمية البيانات التى تتولد يوميًا كناتج للنشاط اليومى لأعمال المؤسسات- أصبح من الصعب إدارة وتنظيم هذه البيانات بالأساليب التقليدية، إلى أن ظهرت تحليلات البيانات الضخمة (Big data analytics)، التى مكنت من إدارتها بكفاءة وفاعلية، فأصبحت مصدرًا أساسيًا لاتخاذ القرارات، إذ يمكن القول أن نجاح القرار وفشله يتوقف على التنظيم الجيد للبيانات، واستثمارها وهذا ما تقوم به تحليلات البيانات الضخمة. (فتوح، 2017)

فالبيانات هي عبارة عن حقائق أولية خام غير منظمة وغير مرتبط بعضها ببعض، ولا تفهم أو يفاد منها إلا بعد معالجتها، فهي- إذا- المعطيات الأولية أو البكر التى تستخلص منها المعلومات، والتى بدورها تعرف بأنها البيانات المعالجة والمنظمة والمترابطة والمعدة للاستخدام واتخاذ القرارات، وهذا يعنى أن المعلومات التى يمكن الإفادة منها فى اتخاذ القرارات وحل المشكلات تستخلص عادة من البيانات.

والبيانات المقصودة هنا هى البيانات التى يتم معالجتها –عادة- بأدوات إدارة قواعد البيانات أو بالطرق التقليدية لمعالجة البيانات كالحواسيب ومحركات البحث؛ لأنها بيانات مهيكلة ومنظمة، وغالبًا ما تكون نوعًا واحدًا، وذات حجم يقاس بوحدات قياس البيانات: كيلوبايت، وميغابايت، وغيغابايت، وإذا كانت البرمجيات التقليدية يمكنها –عادة- التعامل مع مجموعات البيانات ذات أحجام الميغابايت والكيلوبايت، فإن أدوات البيانات الضخمة يجب أن تكون قادرة على التعامل مع مجموعات البيانات ذات أحجام التيرابايت والبيتابايت. (Wang et al., 2016)

وفى الواقع إنه من الصعب تعريف مصطلح (Big Data) بشكل دقيق؛ لاختلاف ما يمكن اعتباره بيانات ضخمة من مؤسسة إلى أخرى ؛ نظرًا لاختلاف قدرات كل مؤسسة على حدة، ولكن بشكل عام عندما نذكر (Big Data)، فنحن نتحدث عن بيانات متعددة الأنواع والمصادر، بحجم يصل إلى المئات من التيرابايت أو البيتابايت أو أكثر، ويمكن تقسيم تعريفات البيانات الضخمة إلى قسمين هما :

التعريف الاصطلاحى

- البيانات (Data) : تمثل الحقائق، وهى الأرقام أو التعليمات المقدمة فى شكل مناسب؛ لفهمها وتفسيرها بما يتناسب مع عمليات الاتصال أو المعالجة عن طريق الإنسان أو الكمبيوتر. (Reitz, 2002)
- البيانات الضخمة: هى كميات من البيانات كبيرة الحجم، عالية السرعة، متنوعة الصيغ، بحيث تحتاج تكلفة باهظة ومجهودًا ضخمًا؛ لإمكانية معالجتها بإستخدام تقنيات قواعد البيانات والبرمجيات التقليدية. (Gartner Glossary)
- البيانات الضخمة: هي عبارة عن مجموعات من البيانات الهائلة التى يمكن تحليلها بواسطة الحاسب الآلى؛ للكشف عن الأنماط والاتجاهات والمؤسسات، وبخاصة فيما يتعلق بسلوك الإنسان وتفاعله ويتجه الكثير من شركات تكنولوجيا المعلومات نحو إدارة وحفظ البيانات الضخمة. (Dictionaries, 2017)
- البيانات الضخمة: هي عبارة عن تراكمات من البيانات الضخمة للغاية والمعقدة لدرجة أنه لا يمكن معالجتها بأدوات قواعد البيانات التقليدية. (Merriam-Webster, 2017)

تعريفات متخصصة

- **البيانات الضخمة:** مجموعة من البيانات لا يمكن الحصول عليها، وتخزينها، وإدارتها بواسطة برمجيات قواعد البيانات التقليدية، وإن حجم هذه البيانات ينمو بمرور الوقت أو بالتقدم التكنولوجى، وتظهر فى تطبيقات يختلف بعضها عن البعض. (Chen et al., 2014)
- **البيانات الضخمة :** جيل جديد من التكنولوجيات صمم لاستخراج قيمة اقتصادية من كميات كبيرة جدًا لمجموعة من البيانات وذلك بالتقاطها واكتشافها وتحليلها. (Gantz & Reinsel, 2011)
- **البيانات الضخمة** لديها القدرة على نمو الاقتصاد العالمى عن طريق تعزيز الإنتاجية، والقدرة التنافسية للمؤسسات. (Chen & Zhang, 2014)
- **البيانات الضخمة:** أصول المعلومات التى تتسم بالحجم العالى، والسرعة والتنوع، وتتطلب تكنولوجيا محددة، وأدوات تحليل تحويلها ونقلها إلى قيمة. (De Mauro et al.., 2016)

- **البيانات الضخمة:** عملية مهتمة بالاستكشاف والتطور، وتطبيقات الخوارزميات، والهياكل الأساسية، وأدوات التنظيم والتكامل والاسترجاع والتحليل والتصور، وهى بيانات ضخمة ومعقدة وغير متجانسة. (Ahmed & Ameen, 2017)
- **البيانات الضخمة:** تنوع كمى ونوعى ضخم من البيانات؛ نتيجة تسارعها وإتاحتها آليًا، ومن ثَم تتطلب معالجة واسعة، والتنقيب عنها بأدوات تكنولوجية غير تقليدية. (العليمى، 2018)
- **البيانات الضخمة:** البيانات التى تتجاوز قدرة معالجة أنظمة قواعد البيانات التقليدية؛ لأنها تتسم بالضخامة، وسرعة الحركة، وتتطلب طرق بديلة لمعالجتها؛ للحصول على أقصى إفادة منها. (Dumbill, 2013) وتوصف بأنها تخزين وتحليل مجموعات من البيانات الكبيرة أو المعقدة باستخدام سلسلة من التقنيات. (Ward & Barker, 2013)
- **البيانات الضخمة:** مجموعة بيانات كبيرة ومعقدة للغاية تتطلب قدرًا من المعالجة السريعة؛ ولذا يصعب التعامل معها باستخدام قواعد البيانات وأدوات التحليل التقليدية. (هيئة تنظيم الاتصالات بسلطنة عمان، 2016)
- يشير كل من **"وانغ واكسو" إلى أن مفهوم البيانات الضخمة** ظهر لأول مرة على يد "لانى" (Wang et al.., 2016). فقد وصف "لانى" البيانات الضخمة بأنها هي البيانات التى لا يمكن معالجتها بواسطة أدوات إدارة البيانات التقليدية ، وأن هذه البيانات لكي تكون ضخمة؛ يجب أن تتصف بثلاث خصائص , تبدأ بحرف " V", وهي : الحجم "volume ", والسرعة " Velocity" , والتنوع "Variety " . (Armour, 2012)
- **البيانات الضخمة:** أرصدة من المعلومات تمتاز بضخامة الحجم والسرعة والتنوع الذى يتطلب أشكالًا مبتكرة وفعالة؛ لمعالجتها بطريقة تختلف عن معالجة البيانات العادية بحيث تمكن مستخدمها من تحسين الرؤية واتخاذ القرارات وعملية التشغيل الآلى. Lehong & (Laney, 2013)

<u>التعريف الإجرائى</u>

البيانات الضخمة هى بيانات ذات حجم كبير ولا يمكن معالجتها بالطرق التقليدية وتتطلب قدرًا من المعالجة السريعة وتقنيات حديثة لمعالجتها وتحليلها.

ولا شك أن مصادر البيانات الضخمة متعددة حيث تشمل خدمات الويب، ووسائل التواصل الاجتماعى، وخدمات البيانات المفتوحة أي: البيانات الحكومية، ومقتنيات المكتبات والمستودعات، والمحفوظات الرقمية، وبيانات الهواتف الذكية، وبطاقات الائتمان، والتلفزيون، والحاسبات، وأجهزة الاستشعار المجهزة، ونظم المراقبة، والأقمار الصناعية، وأنظمة الملاحة. (Shaw, 2014)

إذا فإن البيانات الضخمة بشكل عام هي المواد الخام التى يصبح لها قيمة عندما يضاف لها معنى لتتحول إلى معلومات مفيدة، وهى بذلك تمثل أيضًا المخرجات لعملية معالجة البيانات.

إن قيام الثورة المعلوماتية والتطور في التكنولوجيا التى تتعامل مع البيانات ومعالجتها- أدى ذلك إلى زيادة أحجام البيانات بشكل كبير ومستمر، وأصبح يشار إليها بالبيانات الضخمة التى يقاس سعتها بالتلريون بايت وأكثر. (Mishra, 2015)

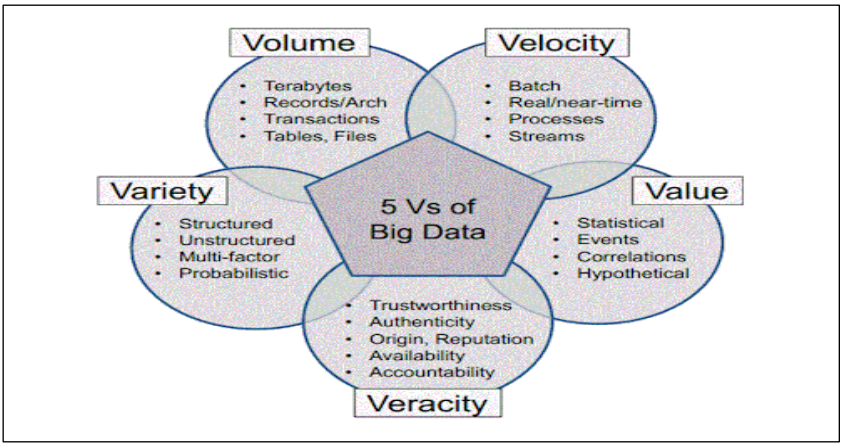
ومما سبق يتضح لنا أن البيانات الضخمة تتسم بما يلي :

- **الحجم :** ويقصد به كمية البيانات المتولدة الذى قد يصل إلى عدد مأهول من البيانات، ويحدد حجمها قيمتها.
- **التنوع :** تشمل البيانات الضخمة بيانات منظمة أو مهيكلة - وتمثل النسبة الأقل- وبيانات غير منظمة أو غير مهيكلة، وتمثل الجزء الأكبر من البيانات، أو خليط بين الاثنين تسمى بالبيانات شبه المنظمة أو المهيكلة.
- **السرعة :** تواتر حدوث البيانات وتتابعه إضافة إلى المعالجة التى تتم لمجموعة من البيانات فى مدة معقولة من الوقت.

وأضافت دراسة "**ميشرا** " (Mishra, 2015) خاصيتين هما :

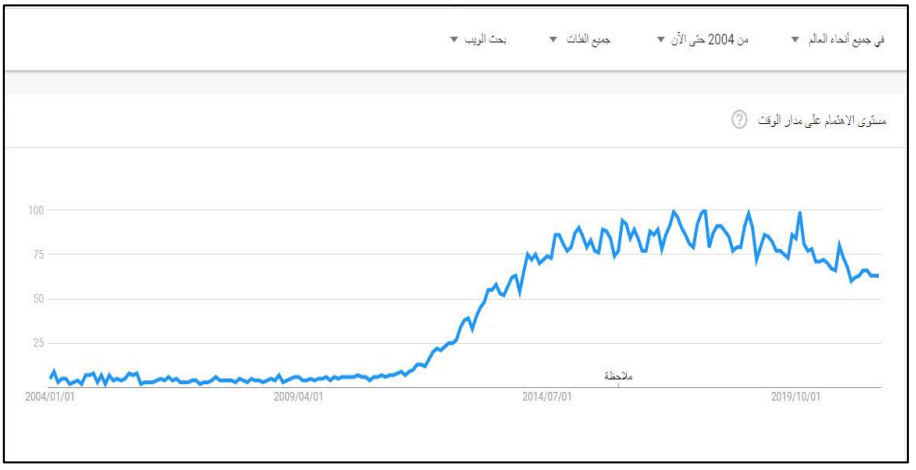
- **التباين :** يقصد به عدم اتساق البيانات فى بعض الأحيان؛ مما يؤدي إلى عدم معالجة البيانات وإدارتها بشكل فعال.
- **الصدق :** يرتبط الصدق بجودة البيانات التى يتم الحصول عليها، ويتطلب ذلك تحليل دقيق لها من حيث فائدتها مع تحرى مصدرها وصحتها.

- وأضاف " **زيكارى**" المزيد من الخصائص وبين التحديات، وقد صنف هذه التحديات والخصائص إلى ثلاثة مجالات: البيانات (Data) ، والعملية (process)، والإدارة (Management). (Zicari, 2014)



شكل رقم (1) خصائص البيانات الضخمة (المصدر:Demchenko et al, 2015)

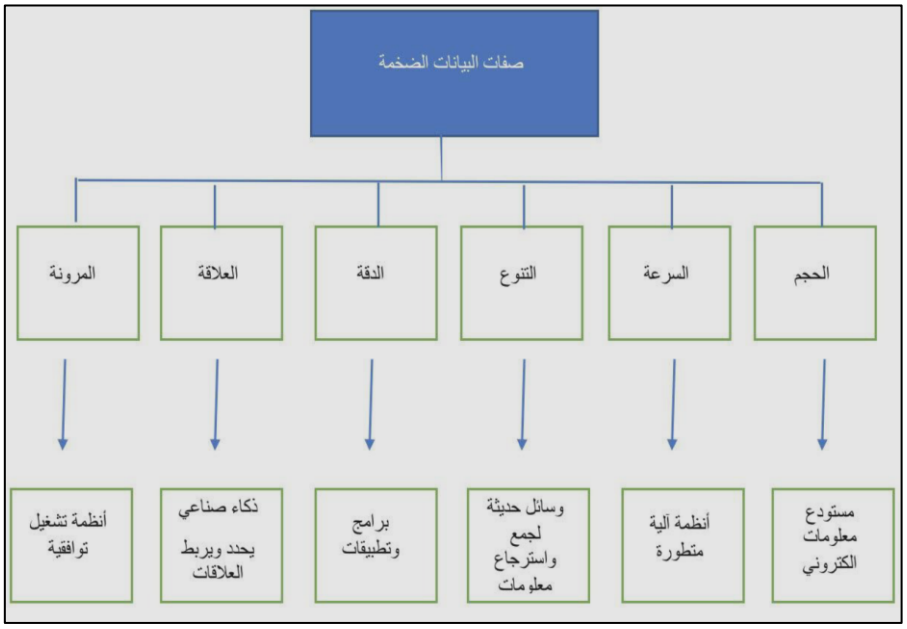
أي أن البيانات الضخمة تتميز بضخامة الحجم ؛ لأنها تهتم بجمع جميع المعلومات الدقيقة؛ لتفسير الأحداث ووضع الحلول.



شكل (2) اتجاهات بحث جوجل حول البيانات الضخمة (GoogleTrends, 2021)

كما أنها تتميز بالسرعة؛ لأنها وليدة اللحظة فتنمو بسرعة فائقة ولأنها تجمع بين طياتها كل معلومة مهما كانت دقيقة، لتفسير الأحداث، واستقراء المستقبل، بالإضافة إلى أن هذه البيانات الضخمة متنوعة؛ لأنها تؤخذ في شتى المجالات والتخصصات سواء أكانت هذه البيانات منظمة أو غير منظمة، مما جعل لها بعدًا فاعلًا فى إثراء المعلومات والمعرفة؛ لمعالجتها للظواهر والأحداث من جوانب عدة، وهذه البيانات الضخمة غاية فى الدقة؛ لأنها تهتم بالتفاصيل الدقيقة وتحاول الاستفادة من أدق المعلومات لتفسير وشرح الظواهر والاحتمالات المتوقع حدوثها على أساس تحليل هذه البيانات وصناعة المعلومات، لتكون نواة للمعرفة المحفز للإبداع والابتكار، ومن مواصفات البيانات الضخمة –أيضًا- أنها تبنى علاقات بين هذه البيانات وتربطها ببيانات أخرى وتحللها؛ لتخلق منها المعلومات والمعرفة التى تشرح أى ظاهرة وتفسرها وتبتكر الحلول لها. (Fister, 2015)

ويوضح الشكل التالى صفات البيانات الضخمة :



شكل رقم (3) صفات وخصائص البيانات الضخمة(المصدر: Fister, 2015)

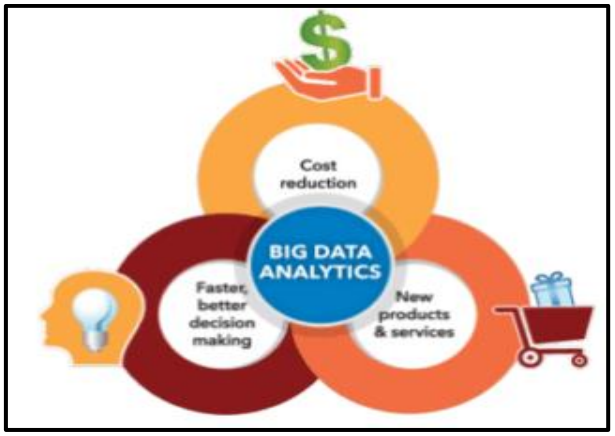
فإذا كانت البيانات الضخمة تتولد باستمرار من خلال تطبيقات مهمة كالهواتف المحمولة، والمواقع الإلكترونية، ووسائل التواصل الاجتماعى وغيرها، فإنها بذلك لا تعد تقنية – في حد ذاتها- ولكنها بيانات كبيرة ومتنوعة وديناميكية، لا يمكن التعامل معها وتنظيمها بالوسائل التقليدية؛ لأنها تستخدم تكنولوجيا حديثة فى التحليل والتخزين التي جعلت من التعامل معها واستخدامها أكثر وصولًا واستفادة ، وأنه - افتراضًا - يمكّن العديد من المؤسسات من اكتشاف المزيد من المعرفة والوصول إليها، وأن استخدام البيانات الكبيرة والاستفادة منها يحدث بشكل يومى فى الأنشطة المختلفة. (Klievink et al., 2017)

ويرجح الاتجاه نحو البيانات الضخمة بسبب المعلومات الإضافية المشتقة من تحليل مجموعة واحدة كبيرة من البيانات ذات الصلة، بالمقارنة مع المجموعات المنفصلة الأصغر حجمًا مع نفس الحجم الإجمالى للبيانات، مما يسمح بوجود ارتباطات. (Kusnetzky, 2017)

والبيانات الضخمة عادة ما تتضمن مجموعة بيانات ذات أحجام تتخطى قدرة البرامج التى يشيع استخدامها؛ لالتقاط وإدارة ومعالجة البيانات فى غضون فترة زمنية مقبولة، وبالنسبة لأحجام البيانات الضخمة فهى هدف متحرك باستمرار، فاعتبارًا من عام 2012م يتراوح حجمها بين بضع عشرات من التيرابايت إلى العديد من البيتابايت من البيانات فى مجموعة واحدة فقط. ومع هذه الصعوبة يتم تطوير منصات جديدة من أدوات " البيانات الضخمة"؛ للتعامل مع مختلف الجوانب الخاصة بالكميات الكبيرة من البيانات.

9/ أهمية البيانات الضخمة

تكتسب البيانات الضخمة أهميتها من كونها تُستخدم على نطاق أوسع فى القطاع المصرفى؛ لفهم العملاء، وفى القطاع التعليمي؛ لتحديد الطلاب المعرضين للخطر، وتمكين الطلاب من تحقيق تقدم دراسى، وفى العمل الحكومى مثل إدارة المرافق، والوكالات، والتعامل مع الازدحام المرورى، أو منع جريمة، وفى القطاع الصحى: مثل إدارة سجلات المرضى، وإعداد الخطط العلاجية والوصفات الطبية، وفى قطاع التصنيع؛ لتحسين الإنتاج والجودة. (Shaw, 2014)



شكل رقم (5) أهمية البيانات الضخمة

ولا شك أنَّ تحليل البيانات الضخمة يوفر إمكانات؛ لإيجاد إجابات للكثير من المشكلات المعقدة التى تواجه الإنسانية، والتى تحتاج أوقاتًا أطول لدراستها، مثل: تغير المناخ، والصراعات، والتنوع البولوجى، والهزات الأرضية، والهجرات. (Mutula, 2016)

كما أنَّ البيانات الضخمة تعمل على إضفاء الطابع الديمقراطى على العمل مع البيانات، وتوفير فرص متعددة؛ لتوليد قيمة ورؤى جديدة من الكميات الضخمة من البيانات. فأما ما يخص الباحثين فقد أدى الإفراج عن بيانات حكومية واسعة النطاق إلى تحقيق اكتشافات علمية جديدة فى مختلف التخصصات. وأما ما يخص الشركات فإن توافر الكميات الكبيرة من البيانات الحكومية وزيادتها يجعل منتجاتها أكثر فعالية مثل: خدمة التنبؤ بأخبار الطقس التجارية التى تعتمد على الوصول إلى بيانات دائرة الأرصاد الوطنية، وأما ما يخص أعضاء المجتمع المحلى، فإن الوصول إلى كميات كبيرة من البيانات الحكومية يتيح فرصًا؛ لخلق طرق جديدة لفهم مجتمعاتهم وتطويرها، وأما ما يخص المؤسسات المجتمعية مثل: المكتبات يمكن لهذا النوع من البيانات أن يقدم نظرة ثاقبة حول احتياجات المجتمع وتكوينه على مستوى لم يكن موجودًا من قبل. (Bertot et al, 2014)

ويؤكد "معهد **ساس**" أن البيانات الضخمة يمكن تحليلها من قبل المنظمات؛ لتحديد الفرص الجديدة، وهذا –بدوره- يؤدى إلى توفير حلول أكثر ذكاءً، ولا سيما فى قطاع الأعمال مثل

تخفيض التكاليف، واستثمار الوقت، وتطوير المنتجات الجديدة، واتخاذ القرارات الذكية، والارتفاع بالأرباح، كما يمكن استخدامها بكفاءة فى مساعدة الشركات؛ لتحديد الأسباب الجذرية للفشل .(Institute)

فالبيانات الضخمة هي أساس لصناعة المعلومة؛ للحصول على المعرفة ومن خلال تطبيقات تعمل على أساس تحليل هذه البيانات؛ للحصول على مؤشرات وتنبؤات تستخدم لإدارة وتطوير المجتمعات واتخاذ قرارات مبنية على المعرفة (Mayer-Schönberger& Cukier, 2013)، كما أنَّ هذه المعلومات والمعرفة المستقاة من البيانات الضخمة - لها دور بارز فى تعزيز وزيادة الاسهامات العلمية فى شتى المجالات. (Williamson, 2017)

ومما يدل على أهمية البيانات الضخمة هو اعتبارها أحد عناصر إدارة المعرفة المهمة، وقد قام كل من **"ميلتون ولامب"** بالتساؤل حول هذه العناصر لإدارة المعرفة فوجدا البيانات الضخمة فى المرتبة الحادية عشر (11)، ولكن مع هذه التطورات التكنولوجية الحديثة والثورة المعلوماتية أصبح للبيانات الضخمة أدوارًا أكبر فى إدارة المعرفة: فتكنولوجيا **"و اتسون"** والذكاء الاصطناعى يعتبران أحد روافد البيانات الضخمة، والإمكانيات التكنولوجية الهائلة المستخدمة فى تحليلها لصناعة المعلومات، التى هى بحاجة لإدارة وتوظيف فعال فى شتى المجالات وتقود للابتكار وللوصول للتميز والتطور وتحقيق رؤية المستقبل وتجاوز صعوباتها وتحدياتها. (Somashekhar et al, 2017)

10/ مصادر البيانات الضخمة

ذكرت اللجنة الاقتصادية الأوروبية (UNECE) التابعة للأمم المتحدة - فى تقرير لها- تصنيفًا لمصادر البيانات الضخمة كالآتى: (UNECE, 2013)

أ- المصادر الناشئة عن إدارة أحد البرامج

هي المصادر سواء أكان البرنامج أو الإدارة لها تبعية حكومية أو لا، مثل: السجلات المصرفية، وسجلات التأمين، وبنوك الطعام، وبرامج التعليم، والبحث العلمى فى الجامعات، والمراكز البحثية، والتعاملات البنكية للمواطنين وغيرها.(Solon, 2014)

ب- المصادر التجارية أو ذات الصلة بالمعاملات الناشئة عن معاملات بين كيانين

من مصادر البيانات الضخمة معاملات البطاقات الائتمانية,والمعاملات التي تجرى عن طريق الإنترنت بوسائل مثل الأجهزة المحمولة، وهذا النوع قد ازدهر ونما بصورة هائلة مع ظهور التجارة الالكترونية في بداية التسعينات التي أدت إلى تضاؤل البعد المكاني، وتخفيض تكلفة الدعاية والنقل، بالإضافة إلى تبادل الآراء وإعطاء التفضيلات وإمكانية التعرف على العميل بنظرة بانورامية نرى بها كل أنشطته وتفضيلاته وتاريخه الشرائي، ولك أن تتخيل أن موقع (Amazon.com) يعالج ملايين العمليات الخلفية كل يوم، فضلا عن استفسارات أكثر من نصف مليون بائع طرف ثالث، وتملك شركة "**أمازون**" أكبر ثلاث قواعد بيانات "**لينوكس**" في العالم وتصل سعتها إلي 7.8 و 18.5 و 24.7 تيرابايت، كما يحتوي الموقع على حوالي (27) مليون مادة معروضة للبيع، وببيع (426) سلعة كل ثانية.(فؤاد، 2013)

ت- مصادر شبكات أجهزة الاستشعار

ومن مصادر البيانات الضخمة-أيضًا- التصوير بالأقمار الصناعية، وأجهزة استشعار الطرق، وأجهزة استشعار المناخ، ورادارات السرعة على الطرق، فعلى سبيل المثال نجد المشروع الفلكي" **مسح سلووان الرقمي للسماء** "Surve Sloan Digital Sky" , وهو مشروع بحثى أمريكي يهدف إلى المسح الفلكى للسماء,باستخدام تليسكوب عملاق متصل بمستشعرات؛ لجمع البيانات عن تكوين النجوم في مجرة "**درب التبانة**" عندما بدأ جمع البيانات الفلكية في عام 2000، فإنه قد تم جمع بيانات في أسابيعه الأولي - القليلة - أكثر مما جمعه في تاريخ علم الفلك بأكمله، ومع استمراره بمعدل 200 جيجا بايت في الليلة، جمع أكثر من 140 تيرابايت من المعلومات.

ث- مصادر أجهزة التتبع

ومن مصادر البيانات الضخمة-أيضًا- تتبع البيانات المستمدة من الهواتف المحمولة والنظام العالمي لتحديد المواقع (GPS) الذي يستخدم في أكثر من جهاز أشهرها الهواتف النقالة، والسيارات حيث يرشد ال (GPS) أكثر من (100) مليون سائق في كل أنحاء العالم يوميًا ويسجل تحركاتهم كما تستخدم شركة (Windermere Real Estate) إشارات (GPS) مجهولة

مما يقرب من (100) مليون سائق؛ لمساعدة مشتري المنازل لتحديد أوقات قيادتهم من وإلى العمل خلال الأوقات المختلفة لليوم، ومثال آخر: "**مصادم الهيدرون العظيم**" "يملك (150) مليون جهاز استشعار تقدم بيانات (40)مليون مرة في الثانية الواحد، وهناك ما يقرب من (600)مليون تصادم في الثانية الواحدة .ولكننا نتعامل فقط مع أقل من (0.001%) من بيانات تيار الاستشعار، فإن تدفق البيانات من جميع تجارب المصادم الأربعة يمثل (25)بيتابايت.

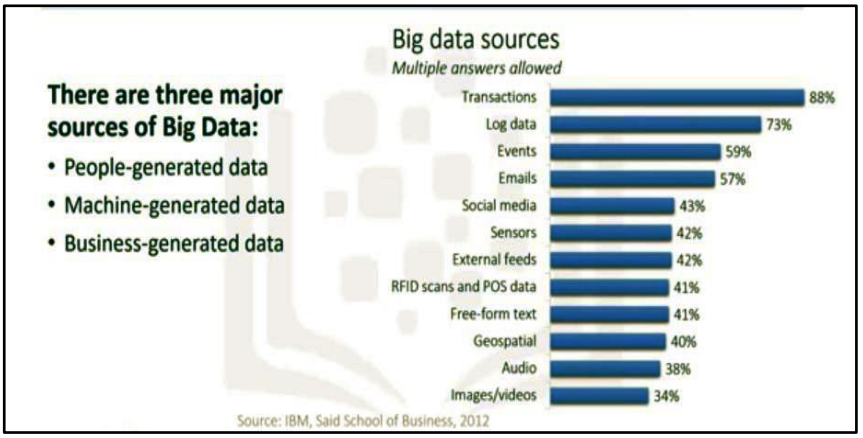
ج- مصادر البيانات السلوكية

ومن مصادر البيانات الضخمة-أيضًا- عدد مرات البحث على الإنترنت عن منتج أو خدمة ما أو أي نوع آخر من المعلومات، ومرات مشاهدة إحدى الصفحات على الإنترنت، فعلى سبيل المثال: (متجر وول) Wal-Mart وهي شركة أمريكية للبيع بالتجزئة بعائدات تبلغ (387.69) مليار دولار أمريكي وتعالج أكثر من مليون معاملة تجارية كل ساعة، والتي يتم إرسالها إلي قواعد بيانات يقدر أنها تحتوي علي أكثر من (2.5) بيتابايت (2560) تيرابايت من البيانات – وهو ما يوازي (167) ضعف البيانات الواردة في جميع الكتب الموجودة في مكتبة (الكونغرس) في الولايات المتحدة. (Banjo, 2014)

ح- مصادر البيانات المتعلقة بالآراء

ومن مصادر البيانات الضخمة-أيضًا- التعليقات على وسائط التواصل الاجتماعي، فقد ذكرت البيانات الرسمية الصادرة عن إدارة **الفيسبوك** - أن قاعدة مستخدمي الموقع حول العالم توسعت لتسجل مع نهاية الربع الرابع من العام 2013 م حوالي (1.86) مليار مستخدم نشط للشبكة الاجتماعية وهي الأكثر شعبية في العالم، منهم حوالي (1.19) مليار مستخدم نشط عبر الأجهزة المتنقلة الذكية. كما يتم رفع أكثر من (300) مليون صورة يوميًا على الموقع بينما تخزن داخل قواعد بياناتها حوالى (50) مليار صورة، واستنادًا إلى آخر البيانات المالية الصادرة عن" إدارة الفيسبوك " العالمية، فإنَّ قاعدة مستخدمي الشبكة زادت بمقدار(165) مليون مستخدم أي: بنسبة (13 %) خلال فترة عام، وذلك مقارنة بعدد مستخدمي الشبكة المسجل في نهاية الربع الرابع من العام(2013 م)، الذي بلغ حينئذ(1.228) مليار مستخدم. (zeforia, 2018)

وقسمت شركة (IBM) مصادر البيانات الضخمة إلى ثلاثة مصادر رئيسة هي : بيانات تُنتج من الإنسان وثانية: تُنتج من التجارة وثالثة: تُنتج من الآلات. (IBM, 2012)



شكل رقم (6) مصادر البيانات الضخمة .المصدر: (IBM, 2012)

11/ تحليلات البيانات الضخمة Big Data Analytics

أكد المجلس الاقتصادى والاجتماعى على أهم الفوائد لتحليل البيانات الضخمة منها: تحسين أداء الخدمات من خلال قبول إعادة استخدام البيانات ثم ربطها بمصادر أخرى؛ مما يؤدى إلى صنع قرارات صائبة؛ وهذا ما يحقق الرضا للمستفيدين من الخدمات المقدمة. (المجلس الاقتصادى والاجتماعى، 2014)

ولتحليلات البيانات الضخمة أهمية كبيرة حيث أنها تقدم ميزة تنافسية عالية للشركات اذا استطاعت الاستفادة منها ومعالجتها؛ لأنها تقدم فهمًا أعمق لعملائها ولمتطلباتهم وهذا يساعد على اتخاذ القرارات المناسبة والملائمة داخل الشركة بطريقة أكثر فعالية بناء على المعلومات المستخرجة من قواعد بيانات العملاء؛ ومن ثَمَّ زيادة الكفاءة والربح وتقليل الخسائر، فباستخدام "وول مارت" لتقنيات وأدوات تحليل البيانات الضخمة استطاعت تحسين نتائج البحث عن منتجاتها عبر الإنترنت بنسبة (10-15%) بينما جاء في تقرير "لماكينزى" - وهى شركة رائدة في مجال استشارات الأعمال- أن القطاع الصحي بالولايات المتحدة لو كان يستخدم تقنيات تحليل البيانات الضخمة بفاعلية وكفاءة لأنتج أكثر من(300) مليون دولار أمريكى فائضًا سنويًا في ميزانية الصحة . (Rouse, 2018)

1/11 تعريف تحليلات البيانات الضخمة

إن تحليلات البيانات الضخمة هي عملية معقدة؛ لفحص مجموعات البيانات الكبيرة والمتنوعة أو البيانات الضخمة؛ للكشف عن المعلومات بما في ذلك الأنماط المخفية والارتباطات غير المعروفة واتجاهات السوق وتفضيلات العملاء التى تساعد على اتخاذ قرارات عمل مستنيرة. (Rouse, 2018)

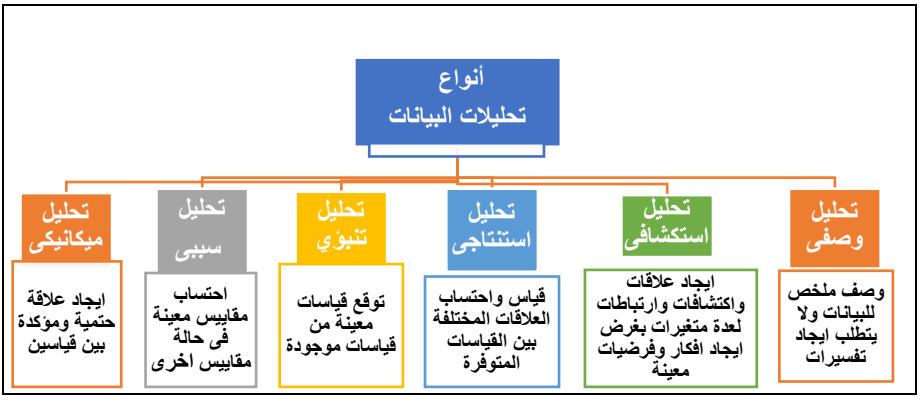
2/11 أهمية تحليل البيانات الضخمة

تتمثل أهمية تحليلات البيانات الضخمة في:

1. رصد واكتشاف أوجه القوة والضعف بناء على نتائج تحليل البيانات.
2. أنها تقدم للمسئولين حلولًا للمشكلات وفقًا لنتائج تحليل البيانات الضخمة بما يتناسب مع طبيعة الجهة أو الهيئة.
3. أنها تزيد من فرص التنافس بما يحققه تحليل هذه البيانات من ميزات ونتائج جديدة تزيد من فرص التطور والتنمية.
4. أنها تساعد على اتخاذ القرارات الصحيحة من خلال نتائج تحليل البيانات، وزيادة فرص التنبوء بالمستويات المستقبلية.

3/11 أنواع تحليلات البيانات الضخمة

تتعدد أنواع تحليلات البيانات الضخمة ولكل نوع أهداف واستخدامات معينة يوضحها الشكل التالى :



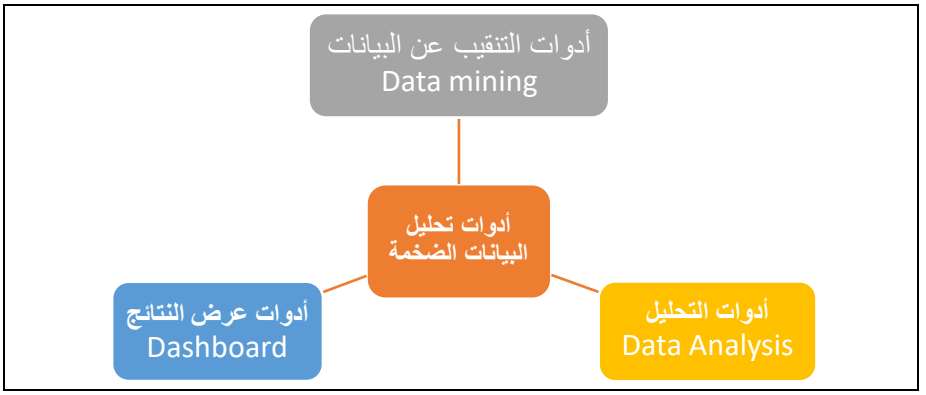
شكل (7) أنواع تحليلات البيانات. (من إعداد الباحث)

تنقسم تحليلات البيانات الضخمة إلى ستة أنواع كما في الشكل هي (Leek, 2015)

1. **التحليل الوصفى** : هو التحليل الذي يستهدف منه عمل ملخص وصفي للبيانات ولا يتطلب إيجاد تفسيرات لها مثل ما يقدمه تحليل البيانات للإحصاء السكانى لدولة معينة، حيث لا يقدم التحليل سوى خلاصة لما يشمله استبيان الاحصاء من جنس وعمر وعنوان وغيره.
2. **التحليل الاستكشافى** : هو التحليل الذي يهدف إلى إيجاد علاقات واكتشافات وارتباطات وميول لعدة متغيرات؛ بغرض إيجاد أفكار وفرضيات معينة، مثل ما قام به مجموعة من الهواة الذين حللوا بيانات فضائية كثيرة جمعها **"مقراب كبلر"** وبعد التحليل وجدوا نظامًا شمسيًا يتكون من أربعة كواكب وذلك بتحليل خصائص الضوء.
3. **التحليل الاستنتاجى**: يعتبر التحليل الاستنتاجي أحد أكثر التحليلات شيوعًا فى البحوث العلمية، ويذهب إلى ما وراء التحليل الاستكشافى ليرى إن كانت الأنماط المكتشفة صالحة لكى تكون وراء مجاميع البيانات المتوفرة، حيث يقوم هذا التحليل بقياس واحتساب العلاقات المختلفة بين القياسات المتوفرة.
4. **التحليل التنبؤى** : يهتم هذا التحليل بتوقع قياسات معينة من قياسات موجودة، مثال ذلك ما تقوم به المؤسسات الإحصائية من تنبؤ نتيجة الانتخابات من خلال تحليل سلوك التنبؤ الذى تتم ملاحظته فى الاستبيانات.
5. **التحليل السببى** : يقوم التحليل السببي باحتساب مقاييس معينة فى حالة تغير مقاييس أخرى، ومثال ذلك احتساب تأثير ممارسة معينة على تقليل الإصابة بمرض معين.
6. **التحليل الميكانيكى** : يقوم هذه التحليل بإيجاد علاقة حتمية ومؤكدة بين قياسين.

4/11 أدوات وبرمجيات تحليل البيانات الضخمة

إن تنفيذ التحليل للبيانات الضخمة يتطلب مجموعة متكاملة من الأدوات والبرمجيات؛ للتحكم والسيطرة على البيانات الضخمة لتحليلها (Warden, 2011) ويمكن تقسيم هذه الأدوات إلى فئات ثلاثة رئيسة هى : (Warden, 2012)



شكل (8) أدوات تحليل البيانات الضخمة (من إعداد الباحث)

1. **أدوات التنقيب عن البيانات Data mining**

تتعامل هذه الأدوات مع البيانات غير المهيكلة كالنصوص وحركات المستخدمين الموزعة على الأجهزة المختلفة عبر الويب.

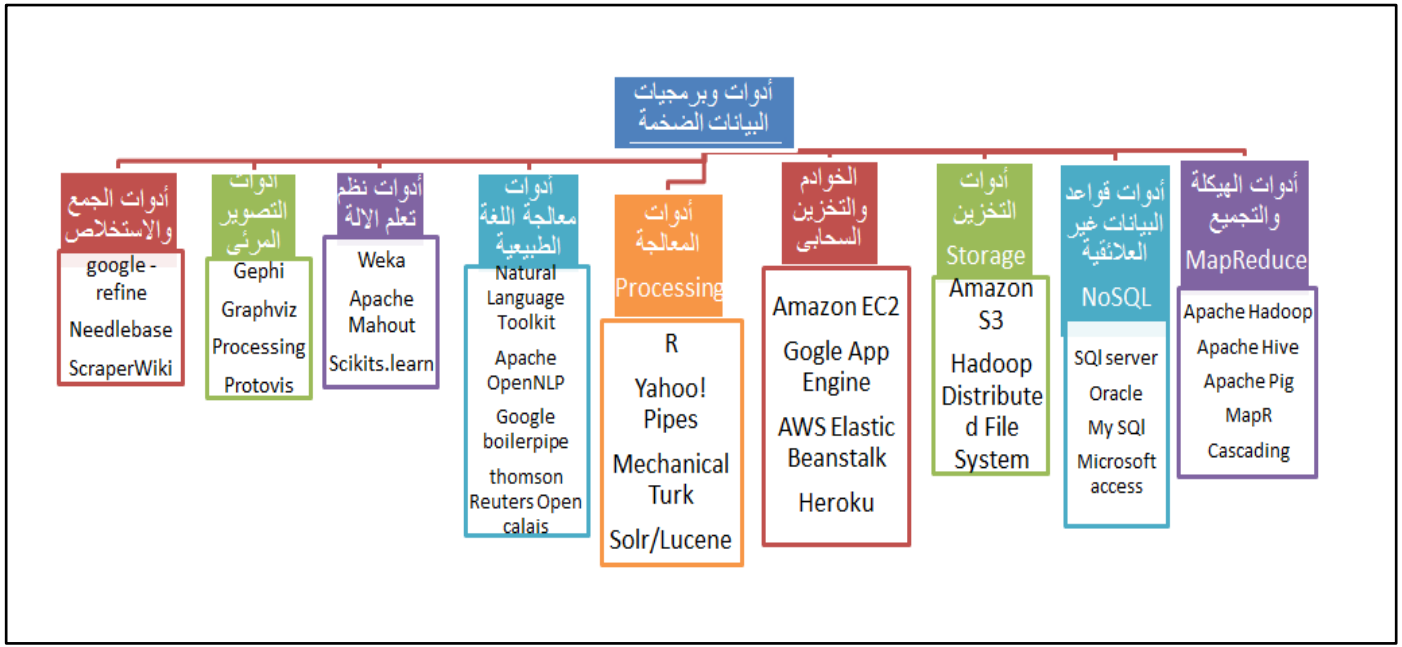
2. **أدوات التحليل Data Analysis**

تستخدم هذه الأدوات للمقارنة والتصنيف والمقاربة والربط؛ للخروج بالنتائج المطلوبة.

3. **أدوات عرض النتائج Dashboard**

تستخدم هذه الأدوات فى عرض النتائج النهائية للتحليل فى شكل مرئى ورسوم بيانية وفقًا لأهداف التحليل التى حددت سابقًا.

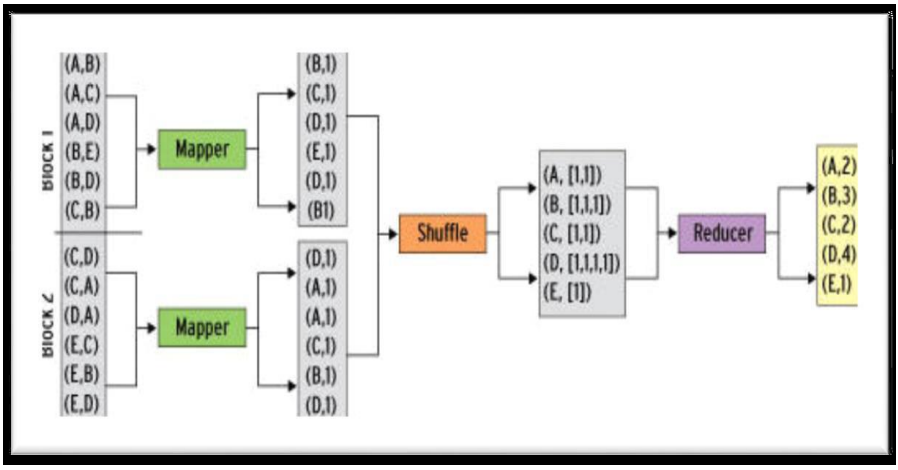
ويمكن تقسيم أدوات وبرمجيات البيانات الضخمة إلى تقسيمات أكثر تفصيلاً كالآتى : (Warden, 2011)



شكل (9) أدوات وبرمجيات البيانات الضخمة (من إعداد الباحث)

1- أدوا ت الهيكلة والتجميع MapReduce

هذه الأدوات عبارة عن نموذج برمجي ونظام تشغيل متكامل للبيانات وتتم عملية الهيكلة والتجميع حيث تتم المعالجة على ثلاث مراحل وهى : التجزئة، والخلط، والتجميع. وتقوم بتجميع البيانات المتشابهة مع بعضها البعض وعمل خريطة لها وترقيمها ترقيمًا تسلسليًا واحدًا لتمييزها، فضلًا عن حذف البيانات المتشابهة. (Dean & Ghemawat.,2004)



شكل (10) مراحل عمل MapReduce. **(المصدر** Dean & Ghemawat.,2004)

2- أدوات قواعد البيانات غير العلائقية NoSQL Databases

هذه الأدوات تخزن البيانات على شكل جداول مرتبطة بعلاقات بينها، وهى قواعد بيانات مفتوحة المصدر وتطلق على قواعد البيانات المتزايدة (Brewer, 2012) . وتتميز هذه الأدوات ببساطة التكلفة وسهولة التنفيذ وإمكانية تجزئة البيانات ونسخها بالإضافة إلى سرعة نقل البيانات وإمكانية التوسع أفقيًا وهذا يعنى سرعة معالجة كمية كبيرة من البيانات.

3- أدوات التخزين Storage

إن من أهم صفات هذه البيانات ضخامة الحجم وبالتالى عند التعامل مع البيانات الضخمة أهم ما يشغل الذهن هو ضمان وسائل تخزين ملائمة بالإضافة إلى كفاءة الاسترجاع، ولذا فقد تم تطوير تقنيات وأدوات تراعى جودة التخزين واسترجاع البيانات الضخمة داخل خوادم ضخمة تستوعب هذه البيانات. (Agrawal & Nyamful, 2016)

4- الخوادم والتخزين السحابى Data Servers& clouding

تعتبر الحوسبة السحابية من أشهر نماذج التخزين على شبكة الإنترنت حيث يتم تخزين البيانات على خوادم ظاهرية متعددة بدلًا من إضافتها على خادم محدد، ومن أهم ميزاتها أنها تتيح مشاركة الإمكانات الحاسوبية والسعات التخزينية وعرض النطاق الترددى، وأثبتت الحوسبة السحابية أنها من الأدوات الأساسية لبناء أنظمة قادرة على التوسع. (Tudoran, 2014)

5- أدوات المعالجة Processing

إن القيمة الأكبر للبيانات الضخمة تكمن فى معالجتها واستخراج قيمتها، وتقسم أدوات المعالجة المشكلات الكبيرة إلى مشكلات صغيرة ويتم حل المشكلات الأصغر فالأصغر، وفى النهاية نجد النتائج المجمعة توفر إجابات نهائية للمشكلة الكبيرة وهذا ما يسمى بمعالجة البيانات على دفعات، وتعتبر برمجيات تحليل البيانات الضخمة هى الأشهر فى عالم البيانات. (Foote, 2016)

6- أدوات معالجة اللغة الطبيعية Natural Language Processing

تعد معالجة اللغة الطبيعية شكل من أشكال الذكاء الاصطناعى وتساعد على قراءة النصوص من خلال محاكاة القدرة البشرية على فهم اللغة، وتمكن هذه اللغة من فهم كلمات منفردة أو مجموعات منها وتساعد أجهزة الكمبيوتر على فهم الجمل كما يتحدث بها الإنسان، وبفضل هذه الأدوات ظهرت تطبيقات جديدة ومفيدة مثل: برامج المساعد للهواتف الذكية، وبرامج الترجمة الفورية، والتعامل البنكى على الخط المباشر، وأدوات التسوق، وبرامج التفاعل

الذكى مع السيارات، وكل هذه البرامج سهلت التواصل مع المؤسسات والشركات، بالإضافة لتحسين الخدمات المقدمة بناء على تحليل أراء المستفيدين.(Team, 2016)

7- أدوات نظم تعلم الآلة Machine Learning systems

إن البيانات الضخمة تتطلب – لتحليلها- أدوات تختلف عن أدوات تحليل البيانات التقليدية لأن كمية العلاقات الناتجة عن البيانات الضخمة تفوق إمكانات أى محلل بيانات مهما بلغت مهارته في جمع الفرضيات واستخراج القيمة منها، ولذا فإنه لا بديل عن استخدام الذكاء الاصطناعى وتعلم الآلة كبديل عن العامل البشري في أداء مثل هذه العلاقات وذلك لأن أدوات نظم تعلم الآلة تتمتع بقدرة كبيرة على استخراج الرؤى والبيانات التى تزداد بصورة كبيرة. (Skytree, 2018)

وتعمل أدوات نظم تعلم الآلة على استخدام البيانات التى تزود بها هذه النظم وتعرف بالبيانات التدريبية والتاريخية، واستخراج العلاقات لتتدرب على أخذ القرار المناسب مع البيانات الشبيهه بها فى المستقبل وتكون المخرجات عبارة عن توصيات أو تجميعات. (Warden, 2011)

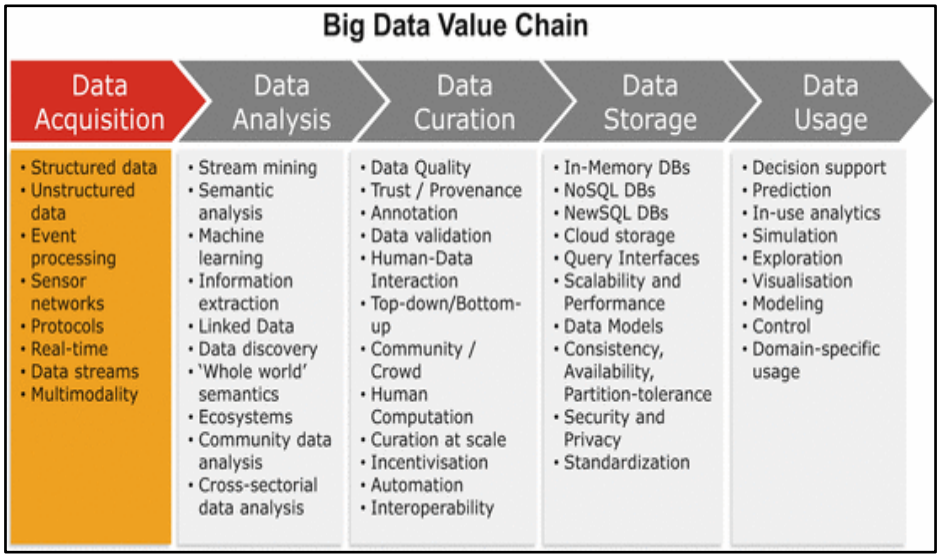
8- أدوات التصوير المرئى Visualization

يعتبر الشق المهم من البيانات الضخمة هو طريقة عرض هذه البيانات بسرعة ووضوح، وبالتالى فإن عرض البيانات بصورة مرئية أو تمثيلها بشكل مرئى هو عبارة عن مجموعة معقدة من البيانات بأسلوب فنى منسق وواضح المعالم على شكل أعمدة مثل: الرسوم البيانية، والخرائط والرسوم التوضيحية، والصور بالإضافة إلى اشتمالها على صورة حية وتفاعلية للبيانات التى يتم جمعها.(Marty, 2015)

9- أدوات الجمع والاستخلاص Acquisition

البيانات الضخمة بها العديد من البيانات ذات القيمة العالية وغالبيتها لا يستحق أن تقوم بعمليات معقدة ومتطورة عليها؛ لأن نتيجتها قد تكون خاطئة ومضللة، بالإضافة إلى أن مجموعات البيانات الضخمة قد تكون سيئة التنظيم ويصعب الوصول إلى المعلومات القيمة التى تحتويها، ولذلك كان لابد من توافر أدوات تعمل على استخلاص البيانات من الشوائب عند

جمعها وقبل وضعها فى مستودعات البيانات الخاصة، وتعتبر هذه العملية من أهم العمليات فى تحليل البيانات الضخمة. (Cavanillas & Wahlster, 2017)



شكل (11) جمع واستخلاص البيانات الضخمة في سلسلة القيمة للبيانات

(المصدر Martin Strohbach,2016**)**

12/ مكونات منظومة البيانات الضخمة

تتكون منظومة البيانات الضخمة من عدة جهات تتفاعل فيما بينها لتحقيق أهداف ورؤية الجهة او الهيئة، هى :(هيئة تنظيم الاتصالات بسلطنة عمان، 2016)

1- **موفر البيانات الضخمة** : يوفر البيانات من مصادرها المختلفة ويقدمها لمقدم الخدمة، وتتمثل أنشطة موفر البيانات فى :

- إنشاء البيانات.
- إنشاء المعلومات الوصفية التى تصف مصدر البيانات والسمات ذات الصلة.
- نشر المعلومات الوصفية.
- إيجاد مصادر البيانات المفتوحة على الإنترنت وإنشاء المعلومات الوصفية لها.

- توفير بيان الخدمات إلى مقدم الخدمة عن البيانات القابلة للاستخدام.

2- **مقدم خدمة البيانات الضخمة** : يقوم مقدم الخدمة بتحليل البيانات الضخمة، وتوفير البنية التحتية اللازمة لها، وتتمثل أنشطته فى :

- البحث فى مصادر البيانات وجمع البيانات عن طريق الطلب المباشر من موفر البيانات أو البحث من خلال الإنترنت.
- تخزين البيانات.
- دمج البيانات.
- توفير أدوات لتحليل البيانات.
- دعم إدارة البيانات مثل: مصدر البيانات، وخصوصية البيانات، وأمن البيانات، وسياسة الاحتفاظ بالبيانات، وملكية البيانات.

3- **عميل خدمة البيانات الضخمة :**

يعتبر العميل هو المستخدم النهائى لمنظومة البيانات الضخمة أو هو نظام يستخدم النتائج أو الخدمات التى يقدمها مقدم الخدمة، كما يمكن للعميل أن ينتج خدمات جديدة اعتمادًا على نتائج تحليل البيانات الضخمة ، وتتمثل أنشطة العميل فى :

- طلب خدمات البيانات الضخمة من المزود.
- استخدام مخرجات خدمات البيانات الضخمة.

13/ البيانات الضخمة فى المكتبات ومؤسسات المعلومات

شكلت البيانات الضخمة مصدرًا مهمًا للبيانات مع انتشار الأجهزة الرقمية مثل الهواتف المحمولة، والهواتف الذكية، والروبورتات الصغيرة من مختلف الأجهزة الرقمية الأخرى، فقد اتسعت أحجام البيانات الضخمة صارت تسير بشكل متوازٍ مع الأجهزة الرقمية الذكية. (Rinne, 2014)

وأكد كل من **"شين ، و انغ، واكسو"** أن خاصية الحجم تنطبق على بيانات المكتبة فى ضوء ارتباطها الوثيق بالبيانات العلائقية التى تشكل شبكة أكبر من البيانات الضخمة، حيث يمكن استخلاص العلاقات الارتباطية من الكتب والمجلات بين المؤلفين المشاركين، والاستشهادات، والأماكن الجغرافية، والتواريخ، والتصنيف الموضوعى، والانتماءات للمؤسسات، والناشرين،

فهذه العلاقة يمكن ربطها فيما بعد بالأعمال، والأحداث، وبراءات الاختراع، ومن ثَمَّ فإن عملية إنشاء وتجهيز وإتاحة هذا الرسم البيانى تسمى بالبيانات الضخمة، ويمكن للمكتبة -أيضا- جمع البيانات التى يبحث عنها المستفيدون، وأما ما يخص السرعة فيمكن العثور على خصائص سرعة البيانات الضخمة فى بيانات المكتبة، إذ أن المكتبة تحتفظ بنسخ متعددة من الملفات على الخوادم فى مواقع موزعة جغرافيًا؛ لذلك هناك تحرك للملفات بين المكتبات وداخلها، بالإضافة إلى الأبحاث الجارية والبيانات البحثية التى تنضم إلى مجموعة البيانات فى المكتبة، وبناءً على ذلك تحتاج بيانات المكتبة إلى معالجة سريعة؛ حتى يتمكن الباحثون من استخدامها، والحصول على نتائج الأبحاث التى يحتاجونها بسرعة ودقة. (Wang et al, 2016)

وفيما يتعلق بالتنوع، فالمكتبة تنتقى أنواعًا مختلفة من البيانات كالكتب، والمجلات والتقارير، والخرائط، والصور، والتسجيلات الصوتية، وبعض البيانات غير المهيكلة والمنظمة، كما تجمع المكتبة مجموعة من البيانات المتعلقة بالاستخدام والمعاملات التى يجريها المستفيدون أثناء تفاعلهم مع أنظمة المكتبات وخدماتها، وتعمل على استنهاض القيمة المحتملة التى يمكن استخلاصها -إلى حد كبير- من البيانات غير المنظمة، لذا فإن سمة التنوع التى تتميز بها البيانات الضخمة يمكن العثور عليها – أيضا- فى بيانات المكتبة. (الشوابكة، 2018)

ولذا فإن بيانات المكتبة تعتبر بيانات ضخمة وذلك بسبب مقتنياتها ذات التنوع الواضح والسرعة العالية والحجم الهائل.

14/ البيانات الضخمة ومجالات الإفادة منها فى المكتبات

تم الربط بين البيانات الضخمة بالمكتبات والتنقيب عنها، ومن أهم مجالات المكتبات التي من تحليل البيانات الضخمة ومعالجتها هى : (سيد، 2016)

1- **إدارة مقتنيات المكتبة** : من خلال تحليل البيانات الضخمة يمكن للمكتبة إدارة مقتنياتها بشكل جديد من خلال جوانب متعددة منها :

- معرفة الثغرات فى مقتنيات المكتبة، حيث تتضح أوجه القصور أو الزيادة فى المقتنيات، مما يساعد متخذي القرار على تقييم مقتنياتهم فى أكثر من جهة، والذي – بدوره- يسهم فى مشاركة المصادر بين المكتبات الأخرى؛ لسد العجز فى المقتنيات.

- إعادة تقسيم مجموعات المكتبة وتكاملها، بما يساعد متخذي القرار فى الحد من تكرار المقتنيات.

2- **قواعد بيانات المستفيدين** : تتوفر بالمكتبة العديد من المعلومات المتعلقة بالمستفيدين، فمن خلال هذه المعلومات تظهر سلوك المستفيدين تجاه مجموعات المكتبة مما يتيح معلومات وبيانات قيمة تساعد فى تطوير خدمات المكتبة وابتكار أشكال جديدة؛ لإتاحة المعلومات بالشكل الذى يرغبه المستفيدون.

3- **تنمية الموارد البشرية** : المكتبة بها متخصصون وغير متخصصين بالإضافة إلى بيانات العاملين وتشمل حالاتهم الاجتماعية وأنواعهم وأعمارهم ومسكنهم وتخصصاتهم ومؤهلاتهم وقدراتهم الاقتصادية وهواياتهم وغير ذلك. ومن خلال تحليل هذه البيانات يمكن استخراج علاقات وروابط قوية، واستنباط أنماط معرفية بما يتيح لمتخذي القرار اكتشاف الموارد البشرية المتاحة لديهم بصورة جديدة ومبتكرة تساعدهم فى إعادة توزيعهم واستخدامهم داخل الأقسام بالمكتبة.

4- **خدمات المعلومات المتاحة بالمكتبة** : ومن خلال إعادة تقسيم المقتنيات وتكاملها وتنمية الموارد البشرية المتاحة بالمكتبة وحسن توزيعها ودراسة سلوك المستفيدين، يصبح أمام المسئولين الرؤية الواضحة لكيفية تقييم الخدمات المقدمة وتحديد مدى جدواها بالإضافة إلى التخطيط الصحيح ؛ لتقديم خدمات جديدة.

5- **الميزانية** : بعد إعادة الهيكلة الجديدة لإدارة المكتبة فمن الطبيعة أن يكون هناك ترشيد فى أوجه صرف الميزانية وتوفيرها بأنماط جديدة لم يمكن التفكير بها من قبل بسبب قصور الميزانية.

وتعد المكتبات واحدة من المؤسسات المناسبة للتعامل مع البيانات الضخمة، حيث الإجراءات التقليدية ولا ينبغى للبيانات الضخمة أن تكون استثناء (Huwe, 2014) ، فالمكتبات تحتاج لفهم التغيرات حتى تتمكن من تحقيق الإفادة القصوى من موظفيها الحاليين من خلال تعديل المسميات وإعادة توزيع الوظائف ذات الصلة، ومن ناحية أخرى حتى تكون قادرة على تحديد المتطلبات الجديدة للمهارات والمعارف الضرورية للموظفين الجدد. (Xia & Wang, 2014)

ومن أهم مشاريع البيانات الضخمة المحتملة التى قد تفيد المكتبات والعاملين بها- دراسات الاستخدام، وتحليل المجموعات، والدراسات المقارنة متعددة التخصصات.(Huwe, 2014)

ويمكن للحكومات – أيضا- تحليل محتوى مواقع التواصل الاجتماعى لمواطنيها حول قرار- أو نظام معين مطبق- تريد تشريعه وتطبيقه لمعرفة ردود الأفعال حول ذلك الأمر من حيث القبول أو الرفض مما يساعد الحكومات فى اتخاذ القرار المناسب لكل حالة.

كما يمكن للشركات - كذلك - الاستفادة من تحليلات البيانات الناتجة من وسائل التواصل الاجتماعى مثل الفيسبوك والتويتر؛ لتحديد جمهورها, والتنبؤ بنتائج حملة التسويق والمبيعات.

كما يمكن للموسيقيين استخدام ملفات مثل سجل الويب (weblog) والبيانات؛ لتحديد تفضيلات الاستماع والتنبؤ بشعبية الأغانى فى المناطق المختلفة مما يساعدهم على تحضير العروض الحية القادمة.

ويمكن التنبؤ بالأحوال الجوية نظرًا للعدد المتزايد من أجهزة الاستشعار التى نملكها اليوم، والملحقة بأجهزة الهواتف الذكية تحديدًا، فإن لدينا بيانات أكثر من أى وقت مضى عن الأحوال الجوية وبدقة عالية، ويمكن للبيانات الضخمة اليوم التقاط صور دقيقة للأحوال الجوية السائدة فى أى مكان فى العالم، والاستفادة من المعلومات المتوافرة فيها لاستخلاص تنبؤات جوية فائقة الدقة.

ويمكن الاستفادة من البيانات الضخمة - أيضًا- فى التنبؤ بالكوارث الطبيعية حيث بات بإمكان مراكز وحدات الاستجابة للكوارث من استخدام البيانات المتعلقة بالجيولوجيا الطبيعية والبيانات الجغرافية للتنبؤ بالكوارث المحتملة من خلال تحليلات البيانات السابقة ومن خلال مقارنة تلك البيانات بالواقع الحالى يتم تعزيز تنبؤات تلك المراكز بهذه النتائج ومن ثَمَّ يتم اتخاذ الإجراءات الاحترازية ووضع استراتيجيات كالإغاثة أو الإخلاء وغيره قبل حدوث الكوارث والأزمات الطبيعية.

ويستفاد بعض مواقع التواصل الاجتماعى مثل الفيسبوك من تحليلات البيانات الضخمة حيث تحدد الفئات التى توجه لها الاعلانات بناء على تحليل اهتمامات المستخدمين وطبيعتهم، وتستخدم تحليلات البيانات الضخمة للمعلومات والبيانات التى يوفرها مستخدمى هذه المواقع

من خلال الرسائل و والتعليقات وحالة المستخدم وميوله واتجاهاته بالإضافة إلى بيانات الملف الشخصى، فكل ذلك يكون عاملًا مهمًا فى توجيه الدعاية والتسويق. (البار، 2017)

15/ أخصائي المكتبات والمعلومات فى بيئة البيانات الضخمة

تزايد حجم البيانات الضخمة بشكل كبير فى العصر الحالى أدى إلى ضرورة وجود متخصصين ومهنيين مؤهلين وقادرين على التعامل مع هذه البيانات الضخمة، فلم يعد أمام أخصائي المعلومات بديل عن تطوير أنفسهم وتطوير مهاراتهم؛ للتعامل مع هذه البيانات الضخمة والاهتمام بما يتعلق بها من آليات تحليل لهذه البيانات وإدارتها وتقديمها؛ لتزيد من فرص الاستفادة منها. (ثابت، 2018)

ولذا أصبح من الضرورى على أخصائى المكتبات والمعلومات إدراك وفهم البيانات الضخمة وكيفية استخدامها والاستفادة منها فضلًا عن إداراك الميزة التنافسية التى ستحققها البيانات الضخمة، مما يحث أخصائى المكتبات والمعلومات على امتلاك مهارات نشر هذه البيانات واستخدامها وتصنيفها وتطوير أساليب استرجاعها للاستفادة منها. (Mutula,2016)

أي أن اختصاصيى المكتبات والمعلومات أصبح عليهم إدراك أدوارهم الجديدة فى علم البيانات، وما يترتب علي من إجادة المهارات الجديدة فى ضوء التطورات التكنولوجية والمعلوماتية الجديدة؛ لتأهيلهم للتعامل مع هذه التغيرات وتلبية احتياجات المستفيدين. (Steyn & de Wee,2017)

كما أصبح من الضرورى – أيضًا- الاعتماد على المهنيين المؤهلين الذين يمتلكون الخبرة في تحليل البيانات؛ لتسهيل استغلالها فى اتخاذ القرارات لأنهم هم الكوادر المستقبلية لعلماء البيانات، ولكن ذلك يتطلب التركيز على مجال تعليم البيانات وإن كان يعانى نقصًا فى المكتبيين الذين يهتمون بهذا المجال. (Stanton, 2018)

ويمكن تصور دور اختصاصيى المكتبات والمعلومات فى عصر البيانات الضخمة كما يلى: (Bertot et al, 2014)

- توفير البيانات التنظيمية والخبرة الإدارية، حيث تتطلب البيانات الضخمة مهارات الإدارة والتنظيم, والحفظ والبيانات الوصفية وهياكل الوصول إليها.

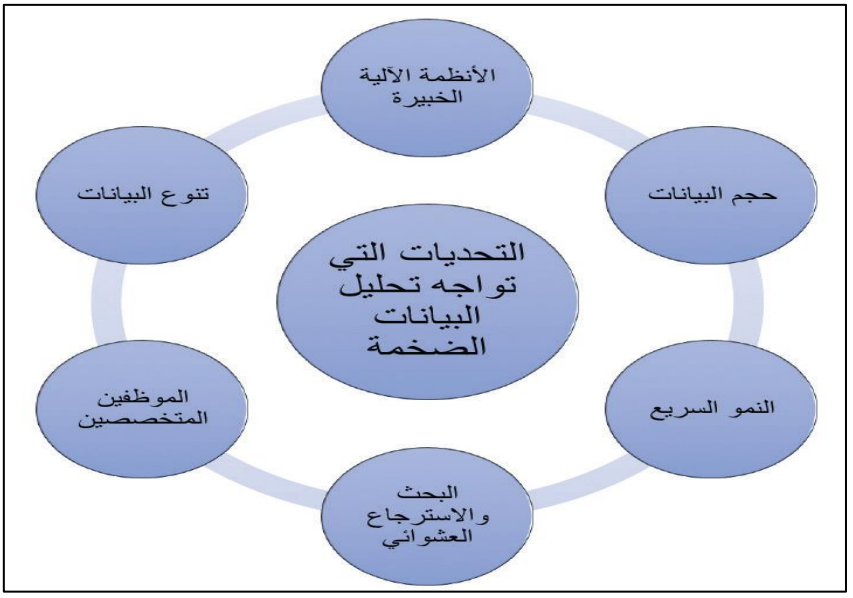
- تطوير مهارات تحليل البيانات داخل المكتبات؛ لتعزيز استخدام البيانات -داخل المجتمعات المحلية - في وضع السياسات واتخاذ القرارات.

ويتمتع أخصائى المكتبات والمعلومات بوضع جيد يمكنهم من مساعدة المستفيدين للتعرف على كيفية العثور والاستفادة من البيانات الضخمة والمحافظة عليها (Hoy, 2014) ، وليس شرطًا أن يكون أخصائى المكتبات والمعلومات على دراية كاملة بالبرمجيات، ولكن لابد أن يكون لديه –على الأقل- بعض الخبرة فيها بالإضافة إلى فهم أساسيات استرجاع المعلومات وأن يكون مدركاً للإحصاءات الاستدلالية الوصفية، وأن يكون ملمًا لكيفية البحث عن أنماط البيانات وعرض البيانات بطريقة فعالة. (Stanton, 2018)

16/ تحديات البيانات الضخمة

لا شك أن المنظمات والهيئات التى تسعى لتطبيق تحليلات البيانات الضخمة تقابلها بعض التحديات والصعوبات أبرزها: (الاكلبى، 2018)

- زيادة حجم البيانات الضخمة بصفة مستمرة.
- النمو الهائل والمتسارع فى كمية البيانات.
- تنوع البيانات.
- الحاجة للموظفين المتخصصين فى تحليل البيانات الضخمة.
- توفير الأنظمة الآلية التى تناسب احتياج الهيئة أو الجهة مع مراعاة تمتع هذه الأنظمة بالمرونة من حيث الاستخدام والتطوير.



شكل رقم (12) تحديات تحليل البيانات الضخمة المصدر : (الاكلبى، 2018).

وهناك بعض التحديات التى تكتنف عملية تنظيم تقديم خدمة تحليلات البيانات الضخمة ومنها : (هيئة تنظيم الاتصالات بسلطنة عمان، 2016)

1- الخصوصية

أصبحت ضمان خصوصية البيانات الشخصية أكثر صعوبة مع تضاعف حجم البيانات وسرعة تبادلها على نطاق واسع، ولذا ينبغى أن تمتثل هذه البيانات لقواعد ومبادئ حماية البيانات وتطبيق ذلك بشكل لا يعوق التقدم لتتلاقى البيانات الضخمة والبيانات الشخصية لتشكل المنتجات الاستهلاكية عبر الإنترنت.

2- أمن البيانات

يتزايد تأمين البيانات مع تزايد سرعة انتقال البيانات وتنوعها وتنوع مصادرها وزيادة حجمها، وبالتالى فإن الآليات الأمنية التقليدية - التى صممت لتأمين البيانات الثابتة فى نطاق محدود - غير كافية لتأمين البيانات الضخمة، وتحليل هذه البيانات الضخمة خارج حدود الدولة المالكة للبيانات يشكل تحدٍ آخر فى تأمين وحماية هذه البيانات.

3- ملكية البيانات

عدم وضوح مالك بيانات القطاع العام أو البيانات الحكومية يشكل تعقيدًا فى عملية مشاركة البيانات الحكومية مع موفر الخدمة، لذا من الضرورى تحديد ما إذا كانت البيانات الحكومية مملوكة للحكومة أو للجهة التى تنتجها، وصلاحيات المالك فى التعامل مع هذه البيانات.

4- تصنيف البيانات الإلكترونية

لا يوجد تصنيف محدد للبيانات الحكومية، وبالتالى تكون هذه البيانات حساسة من حيث سبل التعامل معها سواء من حيث التخزين وقابلية النشر أو المشاركة، بما يشكل تحدٍ فى تقديم خدمة تحليل البيانات الضخمة، وفى ظل غياب هذا التصنيف يتيح لأى جهة أن تتحفظ فى نشر بياناتها الإلكترونية أو مشاركتها مع موفر الخدمة.

5- المسؤولية والمساءلة

هي مسئولية موفر الخدمة فى التعامل مع البيانات، وتتطلب الرقابة الداخلية وضع آليات داخلية وأنظمة تحكم قياسية تضمن الامتثال وتقديم الأدلة بما فيها تقارير التدقيق؛ لإثبات الامتثال للسلطات الرقابية، واستمرارية موفر الخدمة للإمتثال للقانون هو عنصر أساسي من عناصر المسئولية.

17/ نتائج الدراسة

توصلت الدراسة لبعض النتائج ومنها:

تتمثل أهمية تحليلات البيانات الضخمة فى رصد واكتشاف أوجه القوة والضعف بناءً على نتائج تحليل البيانات، وأنها تقدم للمسئولين حلولًا للمشكلات وفقًا لنتائج تحليل البيانات الضخمة بما يتناسب مع طبيعة الجهة أو الهيئة، وأنها تزيد من فرص التنافس بما يحققه تحليل هذه البيانات من ميزات ونتائج جديدة تزيد من فرص التطور والتنمية، و تساعد على اتخاذ القرارات الصحيحة من خلال نتائج تحليل البيانات، وزيادة فرص التنبوء بالمستويات المستقبلية، وتتمثل أدوات تحليل البيانات الضخمة فى ثلاث أدوات رئيسة وهى أدوات التنقيب وأدوات التحليل وأدوات عرض النتائج، وأن بيانات المكتبة تعتبر بيانات ضخمة وذلك بسبب مقتنياتها ذات التنوع الواضح والسرعة العالية والحجم الهائل، ويمكن الاستفادة من تحليلات

البيانات الضخمة فى المكتبة فى إدارة مقتنيات المكتبة وقواعد بيانات المستفيدين وتنمية الموارد البشرية بالاضافة لخدمات المعلومات المتاحة بالمكتبة وفى ترشيد سياسة الانفاق، ويتمثل دور أخصائى المكتبات والمعلومات فى عصر البيانات الضخمة فى توفير البيانات التنظيمية والخبرة الإدارية، حيث تتطلب البيانات الضخمة مهارات الإدارة والتنظيم والحفظ والبيانات الوصفية وهياكل الوصول إليها، وتطوير مهارات تحليل البيانات داخل المكتبات ؛ لتعزيز استخدام البيانات -داخل المجتمعات المحلية في وضع السياسات واتخاذ القرارات.

قائمة المراجع

أبوالنصر، مدحت محمد. (2012). *مقومات التخطيط والتفكير الاستراتيجى المتميز*. المجموعة العربية للتدريب والنشر.

الأكلبى، على بن ذيب. (2018). البيانات الضخمة واتخاذ القرار فى جامعة الملك سعود: دراسة تقييمية لنظام اتقان. *مجلة دراسات المعلومات والتكنولوجيا*، *2* (15)، ص. 2- 12.

البار، عدنان مصطفى. البيانات الضخمة ومجالات تطبيقها. متاح من خلال : https://www.kau.edu.sa/GetFile.aspx?id=285261&fn=Article-of-the-Week-Adnan-Albar-02-November-2017.pdf

ثابت، محمد أحمد. (2018، مارس6-8). *البيانات الضخمة ورهانات الأدوار المتغيرة لأخصائى المعلومات : مقاربة فكرية ورصد تحليلى للكفايات والمواصفات المستقبلية في ضوء تقنيات الويب 2.0* [بحث مقدم]. المؤتمر الرابع والعشرون لجمعية المكتبات المتخصصة فرع الخليج العربى البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفى. مسقط، المغرب.

الرباعى، سليمان بن إبراهيم والطيب، زينب. (2018). الادوار الجديدة لأخصائى المعلومات للتعامل مع البيانات الضخمة. *Journal of Information Studies & Technology*, 16. 1-15.

السالمى، جمال بن مطر بن يوسف وعرابه، سعيد. (2018، مارس6-8). البيانات الضخمة ودورها في دعم اتخاذ القرار والتخطيط الإستراتيجي : دراسة وصفية [بحث مقدم]. المؤتمر الرابع والعشرون لجمعية المكتبات المتخصصة فرع الخليج العربى البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفى. مسقط، المغرب.

سيد، أحمد فايز أحمد. (2016). أدوات التنقيب عن البيانات مفتوحة المصدر: دراسة تحليلية تقييمية. *مجلة جامعة طيبة للاداب والعلوم الانسانية*، *5*(1).

الشوابكة، يونس أحمد إسماعيل. (2018، مارس6-8). *الوعى بمفهوم البيانات الضخمة لدى العاملين فى المكتبات الأكاديمية: دراسة حالة لمكتبة الجامعة الأردنية* [بحث مقدم]. المؤتمر الرابع والعشرون لجمعية المكتبات المتخصصة فرع الخليج العربى البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفى. مسقط، المغرب.

العتيبى، أمال بن سعود. (2018، مارس6-8). *البيانات الضخمة وصناعة المعلومات في المكتبات ومراكز المعلومات : مكتبة الملك فهد الوطنية نموذجاً* [بحث مقدم]. المؤتمر الرابع والعشرون لجمعية المكتبات المتخصصة فرع الخليج العربى البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفى. مسقط، المغرب.

على، احمد خيرى عبدالله. (2018). *دور مراكز المعلومات في إدارة البيانات الضخمة : مؤسسات الرعاية الصحية في مصر نموذجاً* [أطروحة دكتوراه غير منشورة]. جامعة سوهاج.

العليمى، ثروت العليمى المرسى. (2018، مارس6-8). *دور البيانات الحكومية المفتوحة الضخمة فى تمكين التحول نحو الحكومة الذكية: دراسة استكشافية مقارنة* [بحث مقدم]. المؤتمر الرابع والعشرون لجمعية المكتبات المتخصصة فرع الخليج العربى البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفى. مسقط، المغرب.

فتوح، عمرو حسن. (2017). تحليلات البيانات الضخمة ودورها فى دعم إتخاذ القرار فى المكتبات. *مجلة المكتبات والمعلومات العربية*. س37، ع. 4، اكتوبر.

فؤاد، أحمد. (2013). *مبيعات أمازون*. 426 سلعة بالثانية. متاح من خلال سكاى نيوز http://cutt.us/Yzx7y

القصاص، علا نبيل محمد رشاد. (2019). *مشروعات البيانات الرقمية الضخمة في المكتبات : دراسة استكشافية* [رسالة ماجستير غير منشورة]. جامعة المنوفية.

المجلس الاقتصادى والاجتماعى. (2014). البيانات الضخمة وتحديث النظم الاحصائية. متاح من خلال: http://undocs.org/ar/E/CN.3/2014/11

النشرتى، مؤمن سيد (2014). نحو التكامل المعرفى من واقع توظيف الانطولوجيات فى إطار التنقيب عن البيانات: دراسة تحليلية. cybrarian journal، (24)، 162-198.

هيئة تنظيم الاتصالات بسلطنة عمان. (2016). *مشروع البيانات الضخمة (Big Data) : الاطار التنظيمى*. نقلا عن : الاتحاد الدولى للاتصالات، قطاع التقييس، Y.3600، متطلبات وقدرات البيانات الضخمة المعتمدة على الحوسبة السحابية. متاح من خلال : https://goo.gl/aqS9A4.

وزارة الاتصالات القطرية (2014). البيانات الضخمة : تحقيق التوازن بين المزايا والمخاطر. http://www.motc.gov.qa/sites/default/files/lbynt_ldkhm_thqyq_ltwzn_byn_lmzy_wlmkhtr.pdf

اليحيائية، خديجة بنت عبدالله (2018، مارس6-8). البيانات الضخمة : التحديات [بحث مقدم]. المؤتمر الرابع والعشرون لجمعية المكتبات المتخصصة فرع الخليج العربى البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفى. مسقط، المغرب.

Big data analytics. IBM. https://www.ibm.com/analytics/hadoop/big-data-analytics

Big data. (2017). Merriam-Webster. https://www.merriamwebster.com/dictionary/big%20data

Big Data. (2021). GoogleTrends. https://trends.google.com/trends/explore?date=all&q=big%20data

Big Data. Gartner Glossary. https://www.gartner.com/en/information-technology/glossary/big-data

Agrawal, R., & Nyamful, C. (2016). Challenges of big data storage and management. *Global Journal of Information Technology: Emerging Technologies*, *6*(1), 1-10.

Ahmed, W., & Ameen, K. (2017). Defining big data and measuring its associated trends in the field of information and library management. *Library Hi Tech News*.

Al-Barashdi, H., & Al-Karousi, R. (2019). Big Data in academic libraries: literature review and future research directions. *Journal of Information Studies & Technology (JIS&T)*, *2018*(2), 13.

Armour, F. (2012). Introduction to big data, presentation at the symposium Big Data and Business Analytics: Defining a Framework. *Center for IT and Global Economy, Kogod School of Business, American University, Washington, DC*.

Banjo, S. (2014). Wal-Mart Notches Web Win Against Rival Amazon.

https://www.wsj.com/articles/SB10001424052702303417104579544080308837964

Bertot, J., Gorham, U., Jaeger, P., & Sarin, L. (2014). Big data, libraries, and the information policies of the Obama administration. *The Bowker Annual*.

Bertot, J. C., Butler, B. S., & Travis, D. M. (2014). Local big data: The role of libraries in building community data infrastructures. Proceedings of the 15th annual international conference on digital government research,

Brewer, E. (2012). CAP twelve years later: How the" rules" have changed. *Computer*, *45*(2), 23-29.

Cavanillas, J. M., Curry, E., & Wahlster, W. (2016). *New horizons for a data-driven economy: a roadmap for usage and exploitation of big data in Europe*. Springer Nature.

Chen, C. P., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information sciences*, *275*, 314-347.

Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*, *19*(2), 171-209. https://doi.org/10.1007/s11036-013-0489-0

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*, *65*(3), 122-135. https://doi.org/10.1108/LR-06-2015-0061

Dean, J., & Ghemawat, S. (2004). MapReduce: Simplified data processing on large clusters.

Demchenko, Y., Ngo, C., de Laat, C., Membrey, P., & Gordijenko, D. (2013). Big security for big data: Addressing security challenges for the big data infrastructure. Workshop on secure data management,

Dictionaries, O. (2017). *English oxford living dictionaries*. https://en.oxforddictionaries.com/definition/big_data

Dictionary, T. T. T. C. *Data Science* https://techterms.com/definition/data_science

Dumbill, E. (2012). Making Sense of Big Data. *Big Data*, *1*(1), 1-2. https://doi.org/10.1089/big.2012.1503

Fister, B. (2015). Big data or big brother? data, ethics, and academic libraries. *Library Issues: Briefings for Faculty and Administrators*, *35*(4), 293-314.

Foote, K. D. (2016). *Big Data Processing 101: The What, Why, and How*. https://www.dataversity.net/big-data-processing-101/

Gantz, J., & Reinsel, D. (2011). Extracting value from chaos. *IDC iview*, *1142*(2011), 1-12.

Hoy, M. B. (2014). Big data: An introduction for librarians. *Medical reference services quarterly*, *33*(3), 320-326.

Huwe, T. K. (2014). Big data and the library: A natural fit. *Computers in Libraries*, *34*(2), 17-18.

Jharotia, A. K. (2019). Big data technology: Big opportunity for librarians. *GGSIP University*.

Kamupunga, W., & Chunting, Y. (2019). Application of Big Data in Libraries. *International Journal of Computer Applications*, *178*(16), 34-38.

Klievink, B., Romijn, B.-J., Cunningham, S., & de Bruijn, H. (2017). Big data in the public sector: Uncertainties and readiness. *Information systems frontiers*, *19*(2), 267-283.

Koltay, T. (2012). Big data, big literacies? *Information, Communication & Society*, *15*, 663.

Kusnetzky, D. (2017). What is "Big Data?". 2010. *Virtually Speaking. Disponível em:. Acesso em*, *10*.

Leek, J. (2015). The Elements of Data Analytics Style: A guide for people who want to analyze data. In: Leanpub: Victoria, BC.

Lehong, H., & Laney, D. (2013). Toolkit: Board-ready slides on big data trends and opportunities. *Stamford, CT: Gartner*.

Li, J., Lu, M., Dou, G., & Wang, S. (2017). Big data application framework and its feasibility analysis in library. *Information Discovery and Delivery*.

Marty, R. (2015). *Big Data Visualization*. https://www.slideshare.net/zrlram/big-data-visualization-44258309

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.

Mishra, A. S. (2015). Information Professionals and Big Data. *International Journal*, *5*(9).

Mutula, S. (2016). Big Data Industry: Implication for the Library and Information Sciences. *African Journal Of Library, Archives & Information Science*, *26*(2), 93.

Reitz, J. M. (2002). *ODLIS: Online Dictionary of Library and Information Science*. WCSU Libraries.

Rinne, N. (2014). Big Data, Big Libraries, Big Problems?: the 2014 LibTech Anti-talk?

Rouse, M. (2018). *Using big data and Hadoop 2: New version enables new applications*. http://searchcloudcomputing.techtarget.com/definition/Hadoop

Sengupta, S. (2016). Big Data: The Next Big Opportunity For Librarians.

Shaw, J. (2014). Why "Big Data" is a big deal. *Harvard Magazine*, *3*, 30-35.

Skytree, I. (2018). *The Future Is Here – Why Is Machine Learning So Big Right Now?* . https://www.skytree.net/machine-learning/why-do-machine-learning-big-data/

Solon, O. (2014). *A simple guide to Care. data. Wired*. https://www.wired.co.uk/article/a-simple-guide-to-care-data

Somashekhar, S., Sepúlveda, M.-J., Norden, A. D., Rauthan, A., Arun, K., Patil, P., . . . Kumar,

R. C. (2017). Early experience with IBM Watson for Oncology (WFO) cognitive computing system for lung and colorectal cancer treatment. In: American Society of Clinical Oncology.

Sonawane, A. (2018). Big Data: Application in Libraries. *International Journal of Scientific Research in Multidisciplinary Studies*, *4*(1), 22-23.

Stanton, S. (2018). Big Data and the Library Professional. *Journal of the Library Administration and Management Section*, *8*(2), 22-24.

Steyn, C., & de Wee, J. A. (2007). The naked librarian: health librarians in the modern era. *Library Review*.

Stryk, B. (2015). *How do organizations prepare and clean Big Data to achieve better data governance? A Delphi Study* Capella University].

Team, E. a. (2016). *NLP for Big Data: What everyone should know?* . https://www.expert.ai/blog/nlp-big-data-everyone-know/

Tudoran, R. (2014). *High-performance big data management across cloud data centers* ENS Rennes].

Ularu, E. G., Puican, F. C., Apostu, A., & Velicanu, M. (2012). Perspectives on big data and big data analytics. *Database Systems Journal*, *3*(4), 3-14.

UNECE, H. (2013). What Does 'Big Data'Mean for Official Statistics? *Translated by High-level Group on the Modernization of Statistical Production and Services*.

Wang, C., Xu, S., Chen, L., & Chen, X. (2016). Exposing library data with big data technology: A review. 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS),

Ward, J. S., & Barker, A. (2013). Undefined by data: a survey of big data definitions. *arXiv preprint arXiv:1309.5821*.

Warden, P. (2011). *Big data glossary*. O'Reilly Media, Inc.

West, D. M. (2012). Big data for education: Data mining, data analytics, and web dashboards. *Governance studies at Brookings*, *4*(1), 1-10.

Williamson, B. (2017). *Big data in education: The digital future of learning, policy and practice*. Sage.

Xia, J., & Wang, M. (2014). Competencies and responsibilities of social science data librarians: An analysis of job descriptions. *College & research libraries*, *75*(3), 362-388.

Zakir, J., Seymour, T., & Berg, K. (2015). Big Data Analytics. *Issues in Information Systems*, *16*(2).

zeforia. (2018). *The Top 15 Valuable Facebook Statistics – Year in Review*. https://yoursuperinspector.com/wpcontent/uploads/2018/09/Top_20_stats_for_facebook.pdf

Zicari, R. V. (2014). Big data: Challenges and opportunities. *Big data computing*, *564*, 103.