الجمهورية العربية السورية

رئاسة مجلس الوزراء

الجهاز المركزي للرقابة المالية

بحث

تقنيات تحليل ورقابة البيانات الضخمة

دراسة ميدانية – الجهاز المركزي للرقابة المالية – سورية

إعداد الباحث : زياد ابراهيم الخطيب

تقنيات تحليل ورقابة البيانات الضخمة

دراسة ميدانية – الجهاز المركزي للرقابة المالية – سورية

ملخص

هدفت الدراسة إلى التعرف على تقنيات تحليل ورقابة البيانات الضخمة من خلال معرفة ماهية البيانات الضخمة وتقنيات تحليل البيانات الضخمة والبحث في تطبيقات تحليل البيانات الضخمة والذكاء الاصطناعي ونظم وبرامج الإلكترونية لتحليل البيانات و رقابة البيانات الضخمة سواء الرقابة الداخلية أو الرقابة الحكومية وتم اجراء دراسة ميدانية على الجهاز المركزي للرقابة المالية في الجمهورية العربية السورية واستخلاص نتائج الفرضيات المتمثلة بالإجابة على التساؤلات المطروحة من قبل الباحث مثل مدى توافر معرفة جيدة بماهية البيانات الضخمة ومدى استخدم الجهاز الاعلى للرقابة تحليل البيانات اثناء عملية التدقيق. وهل يتوفر لدى الجهاز المركزي الكفاية والملائمة لتحليل البيانات الضخمة. وما مدى تطبيق تقنيات تحليل البيانات الضخمة لدى أجهزة الرقابة (الداخلية والحكومية) والاستفادة منها في خلق قيمة مضافة وجودة وموثوقية للعملية الرقابية عبر مراحلها المتعددة. وتمت الدراسة من خلال المنهج الوصفي التحليلي وعن طريق الاستبانة بهدف التعرف على تقنيات تحليل ورقابة البيانات الضخمة من خلال أنواعها وخصائصها، وتسليط الضوء على الطرق والأدوات المستخدمة لتحليل البيانات الضخمة، وأهم القضايا المتعلقة بها. وتم استخلاص النتائج والتوصيات التي تؤكد على الدور الهام لتحليل البيانات وتقنيات الرقابة والذكاء الاصطناعي والتعلم الآلي على تحليل كمياتٍ كبيرةٍ من البيانات بالاعتماد على الخوارزميات المتقدمة وجب التعرف عليها وضرورة دعم الحكومات على تطبيق تقنيات الذكاء الاصطناعي والتعلم الالي واستخدامها عند التحول الرقمي. وأن تطبيق كافي وملائم لتحليل البيانات الضخمة في القطاع العام يساعد الجهاز الاعلى للرقابة على تحليل البيانات اثناء عملية التدقيق وضرورة توافر تقنيات تحليل البيانات الضخمة لدى أجهزة الرقابة ومستلزمات رقابة البيانات الضخمة من (الأفراد، العمليات ،التكنولوجيا) للقيام بعمليات تحليل البيانات الضخمة والاستخدام الاستراتيجي للبيانات الضخمة في اتخاذ القرارات الرئيسية وتقديم الخدمات.

الكلمات الافتتاحية:

البيانات الضخمة – تقنيات التحليل – تطبيقات تحليل البيانات الضخمة –الذكاء الاصطناعي – نظم وبرامج الإلكترونية لتحليل البيانات – الرقابة الداخلية – الرقابة الحكومية

المقدمة:

في زمن التكنولوجيا الرقمية المتطورة والسريعة، أصبحت البيانات محيطةً بنا من كُل جانبٍ وموجودةً في كُل مكانٍ من حولنا، وذلك نتيجة تزايد استخدام الإنترنت والتكنولوجيا بشكلٍ كبيرٍ وهائلٍ في العالم أكمله، ويتزايد حجم البيانات الرقميّة الموجودة في عالمنا ويتزايد تأثيرها في تغيير حياتنا بشكلٍ سريعٍ جدًا ويتضاعف هذا الحجم كل عامين، حيث تقول دراساتٌ قامت بها فوركس بأن سيتم إنشاء ورفع حوالي 1.7 ميغابايت من المعلومات الجديدة في كل ثانية تمر من حياة كل إنسانٍ على وجه الأرض، مما يجعل مجال البيانات يتطور بشكلٍ كبيرٍ ويرتفع شأنه وأهميته في مستقبل عالمنا وحياتنا، واليوم سوف نستطلع على من علوم البيانات وكذلك تحليل البيانات الضخمة وغيرها من الأمور التي تتعلق بهذا الموضوع.

لقد أطلق صندوق النقد الدولي (IMF)[1] عبارة "البيانات هي النفط الجديد"، فماذا يقصد بالبيانات وهل تستطيع إحداث ثورة مثلما فعل النفط؟

مشكلة البحث:

تحليل البيانات الضخمة يعني استخدام التقنيات التحليليّة المُتقدمة لتحليل وفهم مجموعات ضخمة وكبيرة من البيانات التي تختلف بأشكالها وأنواعها، ويشمل هذا العلم أو المجال بالغ الأهمية تحليل البيانات المُنظمة، وشبه المُنظمة، وغير المُنظمة القادمة من مُختلف المصادر وبمُختلف الأحجام ابتداءً من التيرابايت وصولًا إلى الزيتابايت، حيث تُتيح تلك العملية الفرصة للباحثين والمدققين والاطراف ذات العلاقة والمنظمات لاتخاذ القرارات بشكلٍ أفضل وأكثر سرعة،

وموضوع البيانات الضخمة من حيث المعرفة والفهم لطبيعتها وبنيتها و مصادرها وضرورة التعامل معها بتقنيات تحليل وبرامج ملائمة وكافية وقادرة على مواكبة سرعتها حيث أشارت الدراسة إلى أن البيانات ترد إلى كافة المنظمات في العالم بشكل أسرع مما كانت تتوقع، وخاصة مع اتساع الأعمال التجارية، والتطور التقني والتكنولوجي، أضف إلى حجمها الهائل وتنوعها لذا تكمن مشكلة البحث في معرفة تقنيات التحليل الاحصائية والبرمجية المستخدمة لمعالجة البيانات الضخمة سواء لدى الحكومة أو الاجهزة الرقابية و كفايتها وملائمتها . وتحديد قدرة الأجهزة الرقابية (الداخلية و الحكومية) على تطبيق تقنيات تحليل البيانات الضخمة والاستفادة منها في العملية الرقابية.

[1] - الامم المتحدة، المجلس الاقتصادي والاجتماعي، اللجنة الاحصائية، البيانات الضخمة وتحديث النظم الاحصائية، مارس 2014 ،

فرضيات البحث:

لدينا الكثير من التساؤلات في دراسة موضوع تقنيات تحليل ورقابة البيانات الضخمة نذكر منها :

- هل يتوافر لدى المدققين والاجهزة الحكومية معرفة جيدة بماهية البيانات الضخمة؟
- كيف يمكن قياس وتقييم الاداء بظل عمل الحكومة الإلكترونية و التحول الرقمى ؟
- هل يوجد تطبيق كافي وملائم لتحليل البيانات الضخمة في القطاع العام؟
- هل يتوافر لدى الجهاز الاعلى للرقابة استخدام تفصيلي للبيانات.
- هل يستخدم الجهاز الاعلى للرقابة تحليل البيانات اثناء عملية التدقيق.
- هل يتوافر لدى الجهاز العناصر التالية (الأفراد، العمليات ،التكنولوجيا) للقيام بعملية التحليل.
- هل " القيمة المضافة "هي الاساس في تحليل البيانات الضخمة للأجهزة العليا للرقابة .
- هل تحليل البيانات من قبل الأجهزة العليا للرقابة تضيف قيمة للتقارير والمواطنين.
- هل يوجد استخدام استراتيجي للبيانات في اتخاذ القرارات الرئيسية وتقديم الخدمات.
- هل كامل العمليات الحكومية الرئيسية مؤتمتة.

ومن هذه التساؤلات نقدم فرضيات البحث التالية :

- الفرضية الأولى: يتوافر لدى المدققين والاجهزة الحكومية معرفة جيدة بماهية البيانات الضخمة .

- الفرضية الثانية: يستخدم الجهاز الاعلى للرقابة تحليل البيانات اثناء عملية التدقيق.

- الفرضية الثالثة: يتوفر لدى الجهاز المركزي الكفاية والملائمة لتحليل البيانات الضخمة.

- الفرضية الرابعة: ما مدى تطبيق تقنيات تحليل البيانات الضخمة لدى أجهزة الرقابة (الداخلية والحكومية) والاستفادة منها في خلق قيمة مضافة وجودة وموثوقية للعملية الرقابية عبر مراحلها المتعددة.

ونأمل أن تقدم الدراسة إجابة لهذه التساؤلات من خلال الفصول والبحوث المعدة وكذلك من خلال الدراسة الميدانية التي سوف تؤكد صحة الفرضيات من عدمها والوصول إلى نتائج وتوصيات مفيدة .

أهمية البحث:

البيانات الضخمة big data أصبحت واقع نعيشه، حتى أن قاموس أوكسفورد اعتمد المصطلح و أضافه للقاموس مع مصطلحات مستحدثة أخرى مثل التغريدة tweet.

وإن البيانات الضخمة سببت متاعب كبيرة للبشر لصعوبة الاستفادة منها وعدم القدرة على التحكم فيها بوضعها الذي هي عليه، واستنتجت الدراسة أن البيانات الضخمة تحوي على الكثير من الكنوز والمنافع الغير مستغلة والتي لم يتم الكشف عنها عبر تحليلها واستخلاص النتائج منها فالبيانات الضخمة تختزن الكثير من المنافع والمعارف القيمة المحتملة والمفيدة، وهدفت الدراسة للتعريف بالبيانات الضخمة، وأهم التطبيقات المستخدمة في تخزينها بما يساهم في التعريف بتلك التطبيقات لنشر ثقافة العمل على تحليل البيانات الضخمة، كم بينت الدراسة أهم الفرص والتحديات عندما تتم معالجة البيانات الضخمة وآليات وطرق التعامل مع مشاكل البيانات والمنهجيات الأساسية للتعامل مع طوفان البيانات، مثل الحوسبة الحبيبية، والحوسبة السحابية، والحوسبة المستوحاة من الحيوية التفاعلية.

أمثلة عملية:

– موقع Amazon.com يعالج ملايين العمليات الخلفية كل يوم، فضلاً عن استفسارات من أكثر من نصف مليون بائع طرف ثالث. وتعتمد أمازون علي نظام اللينوكس بشكل أساسي ليتمكن من التعامل مع هذا الكم الهائل من البيانات، و تملك أمازون أكبر 3 قواعد بيانات لينوكس في العالم والتي تصل سعتها إلي 7.8، 18.5 و 24.7 تيرابايت.

– سلسلة المتاجر Walmart تعالج أكثر من مليون معاملة تجارية كل ساعة، والتي يتم استيرادها إلي قواعد بيانات يُقدر 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.

– يعالج فيس بوك 50 مليار صورة من قاعدة مستخدميه. ويقوم نظام حماية بطاقات الائتمان من الاحتيال " FICO Falcon Credit Card Fraud Detection System" بحماية 2.1 مليار حساب نشط في جميع أنحاء العالم.

هدف البحث:

التعريف بأفضل ممارسات عن معالجة البيانات الضخمة التي تم الاطلاع عليها وآليات تحليلها واستخراج المنافع منها، والتعرف على كيفية قيام المنظمات بإعداد البيانات وتنظيمها ليتم استخدامها بشكل مفيد في إدارة الجهات والمنظمات الحكومية من خلال مساعدتها على استخراج معلومات أكثر فعالية من البيانات الضخمة والحصول على قيمة مضافة لها، وبينت الدراسة أن البيانات الضخمة تمثل مشكلة للعديد من المنظمات والحكومات وأجهزة الرقابة في العديد من المجالات والتخصصات وذلك لأن تقنية معالجة البيانات الحالية لا تساعد على جعل البيانات الضخمة بيانات فعالة ومفيدة، وبينت الدراسة أن السبب الرئيسي الذي أدى بالمنظمات للقيام بمعالجة البيانات الضخمة وتحليلها والاستثمار في ذلك كان لأجل تقليل المخاطر التنظيمية والتقنية وتحقيق أرباح أعلى من هذه البيانات، كما توصلت الدراسة إلى أن تنظيم البيانات وتحليلها يساعد في تقديم معلومات مفيدة تساند المسؤولين في اتخاذ القرار وضمان مستقبل قوي للمنظمة، وفي ما يتعلق مدى الإفادة من قدرات الحاسب الآلي في التعامل مع البيانات الضخمة

منهج البحث:

تستخدم هذه الدراسة المنهج الوصفي التحليلي، وتستمد معلوماتها من مصدرين متكاملين هما:

أ. مصادر ثانوية (secondary sources): وتتمثل في مختلف المراجع والمقتنيات المكتبية، بما في ذلك الكتب والدوريات والمواقع الملائمة على شبكة الانترنت.

ب. مصادر أولية (primary sources): والتي تتمثل في الأشخاص المتخصصين الذين تتوافر لديهم المعلومات من خلال ممارساتهم وملاحظاتهم وخبراتهم وتم تجميع المعلومات من هؤلاء عن طريقة الاستبانة

عرض وتحليل الدراسات السابقة:

وتتفق الدراسات على أمور من أهمها أهمية استثمار البيانات الضخمة بعد تنظيمها وتحليلها لمساندة متخذ القرار، وأن ذلك يتم عند توفر الكفاءات البشرية المؤهلة، والأنظمة الخبيرة لتكتمل عملية تحليل البيانات الضخمة بطريقة مثالية ولكي يمكن تقديم مخرجات صحيحة، مع الأخذ في الحسبان مسألة تخزين البيانات الضخمة من خلال مستودعات الحوسبة السحابية التي توفر مساحات تخزين كبيرة مع أهمية ان يتوفر الحد

الكافي من الإجراءات الأمنية المناسبة التي تضمن المحافظة على خصوصية البيانات لتكون بذلك مصدر ثراء وقوة.

- دراسة ستراييك (stryk 2015)[2] فقد تناولت موضوع البيانات الضخمة من جهة التعريف بأفضل ممارسات المسؤولين عن معالجة البيانات الضخمة التي تم الاطلاع عليها وآليات تحليلها واستخراج المنافع منها، وهي من الدراسات التنبؤية الاستشرافية حيث اعتمدت على منهج دلفاي وهو منهج تفاعلي يعتمد على التواصل مع مجموعة من الخبراء أو لجنة من الخبراء من أجل استطلاع آراءهم حول موضوع معين في عدة جولات بواسطة أداة الاستبانة في الغالب، وهو ما تم بالفعل في هذه الدراسة التي استخدمت هذا المنهج للتعرف على كيفية قيام المنظمات بإعداد البيانات وتنظيمها ليتم استخدامها بشكل مفيد في إدارة المنظمة من خلال مساعدة المنظمات على استخراج معلومات أكثر فعالية من البيانات الضخمة والحصول على قيمة مضافة لها، وبينت الدراسة أن البيانات الضخمة تمثل مشكلة للعديد من المنظمات في العديد من المجالات والتخصصات وذلك لأن تقنية معالجة البيانات الحالية لا تساعد على جعل البيانات الضخمة بيانات فعالة ومفيدة، وبينت الدراسة أن السبب الرئيسي الذي أدى بالمنظمات للقيام بمعالجة البيانات الضخمة وتحليلها والاستثمار في ذلك كان لأجل تقليل المخاطر التنظيمية والتقنية وتحقيق أرباح أعلى من هذه البيانات، كما توصلت الدراسة إلى أن تنظيم البيانات وتحليلها يساعد في تقديم معلومات مفيدة تساند المسؤولين في اتخاذ القرار وضمان مستقبل قوي للمنظمة، وفي ما يتعلق مدى الإفادة من قدرات الحاسب الآلي في التعامل مع البيانات الضخمة.

- دراسة بيروقل (Bieraugel 2016) التي تناولت موضوع البيانات الضخمة وضرورة التعامل معها والاستعداد المستمر لمقابلة تضخمها المتسارع، حيث أشارت الدراسة إلى أن البيانات ترد إلى كافة المنظمات في العالم بشكل أسرع مما كانت تتوقع، وخاصة مع اتساع الأعمال التجارية، والتطور التقني والتكنولوجي، والتوسع في التعليم العام والعالي، وهذا كله يولد بيانات تزداد تضخما ما جعلها مثار نقاشات وتكهنات لدى المسؤولين والخبراء بسبب الخوف من عدم القدرة على استيعابها أو السيطرة عليها، وأكدت الدراسة على أنه من المهم أن نفهم أن البيانات الضخمة يمكن أن تكون مصدر ممتاز وثري للمعلومات الموجودة في المكتبات عندما يتم التعامل مع هذه البيانات من خلال تحليلها واستخدام الأدوات الخاصة بذلك والتي ستؤثر على المكتبات بشكل مباشر من خلال معالجة ما لديها من بيانات ضخمة عن المقتنيات

[2] Stryk B. 2015;, do organizations prepare and clean Big Data to achieve better data governance? A Delphi Study (doctoral dissertation, Capella University

والمستفيدين والخدمات التي تقدمها وعمليات الإعارة والنسخ ، وبشكل غير مباشر من خلال التعاون مع أعضاء هيئة التدريس من خلال تحليل البيانات الواردة في نتائج أبحاثهم والتوصيات التي يتوصلون لها وتحويلها إلى معلومات مفيدة وقيمة مضافة لمقتنيات المكتبات.

– دراسة جانسن وآخرون (Janssen 2017)[3] بإجراء دراسة حول تأثير البيانات الضخمة على اتخاذ القرار حيث أظهر هذا البحث الذي هو عبارة عن دراسة حالة أن قناعة المنظمات بأهمية الاستفادة من البيانات الضخمة هي عملية تطورية يؤدي فيها التفهم التدريجي لإمكانات البيانات الضخمة والقيمة المضافة لها.

خطة البحث:

مسلسل	موضوع	الصفحة
الفصل الاول	البيانات الضخمة	8
البحث الاول	مقدمة	8
البحث الثاني	ماهية البيانات الضخمة	9
البحث الثالث	تقنيات تحليل البيانات الضخمة	16
البحث الرابع	تطبيقات تحليل البيانات الضخمة	20
البحث الخامس	الذكاء الاصطناعي وعلم البيانات	24
البحث السادس	نظم وبرامج الإلكترونية لتحليل البيانات	26
الفصل الثاني	رقابة البيانات الضخمة	32
البحث الاول	الرقابة الداخلية والبيانات الضخمة	32
البحث الثاني	الرقابة الحكومية للبيانات الضخمة	37
الفصل الثالث	دراسة ميدانية – الجهاز المركزي للرقابة المالية – سورية –نتائج – توصيات	41
	المراجع العربية والاجنبية	52

[3] Janssen M, van der Voort H, Wahyudi A. 2017;, Factors influencing big data decision-making quality. Journal of Business Research, 70 (2017) 338-345. DOI: https://doi.org/10.1016/j.jbusres.2016.08.007 .

الفصل الاول :البيانات الضخمة

البحث الاول : مقدمة

البيانات الضخمة مصطلح يشير إلى مجموعة بيانات تستعصي لضخامتها أو تعقيدها على التخزين أو المعالجة بإحدى الأدوات أو التطبيقات المعتادة لإدارة البيانات. أو ببساطة لتقريب الأفهام، لا يُمكن التعامل معها على حاسوب عادي بمفرده من خلال قاعدة بيانات بسيطة. ومن سمات مجال «البيانات الضخمة» استعمال حواسب عديدة لتقاسم الأعمال المطلوبة

وهي عادة ما تتضمن مجموعات بيانات ذات أحجام تتخطي قدرة البرامج التي يشيع استخدامها لالتقاط وإدارة ومعالجة البيانات في غضون فترة زمنية مقبولة. وبالنسبة لأحجام البيانات الضخمة فهي هدف متحرك باستمرار، فاعتباراً من عام 2012، يتراوح حجمها بين بضع عشرات من تيرابايت إلي العديد من بيتابايت من البيانات في مجموعة واحدة فقط. ومع هذه الصعوبة، يتم تطوير منصات جديد من أدوات "البيانات الضخمة" للتعامل مع مختلف الجوانب الخاصة بالكميات الكبيرة من البيانات.

في تقرير بحثي وعدد من المحاضرات المتعلقة به عام 2001، قام "دوغ لاني" محلل مجموعة META Group (المعروفة الآن باسم Gartner)[4] بتعريف تحديات نمو البيانات وفرصها كعنصر ثلاثي الأبعاد، بمعني زيادة الحجم (كمية البيانات)، السرعة (سرعة البيانات الصادرة والواردة) والتنوع (تنوع أنواع البيانات ومصادرها). وتقوم Gartner والكثير من الشركات في هذه الصناعة الآن بالاستمرار في استخدام نموذج "Vs3" لوصف البيانات الضخمة. وفي 2012، قامت Gartner بتحديث تعريفها ليصبح كالتالي: "البيانات الضخمة هي أصول معلومات كبيرة الحجم، عالية السرعة، وعالية التنوع تتطلب أشكال جديدة من المعالجة لتعزيز عملية صنع القرار والفهم العميق وتحسين العملية".

البحث الثاني : ماهية البيانات الضخمة

1-تعاريف :

تعريف TBDI للبيانات الضخمة: البيانات الضخمة هو مصطلح ينطبق علي الأجسام الضخمة للبيانات التي تتنوع في طبيعتها سواء أكانت منظمة، غير منظمة أو شبه منظمة، بما في ذلك من المصادر الداخلية

[4] موقع سين نيوز - ما هي تحليلات البيانات الكبيرة؟ 2021

أو الخارجية للمنظمة، ويتم توليدها بدرجة عالية من السرعة مع نموذج مضطرب، والتي لا تتفق تماماً مع مخازن البيانات التقليدية والمنظمة وتتطلب نظام إيكولوجي قوي ومعقد مع منصة حوسبة عالية الأداء وقدرات تحليلية للالتقاط ومعالجة وتحويل وكشف واستخلاص القيمة والرؤى العميقة في غضون وقت زمني مقبول

البيانات الضخمة هي تلك البيانات التي تمتلكها جهة ما بكميات كبيرة لتستعملها في مجال واحد لا يوازي تكلفة جمعها وتخزينها.

البيانات الضخمة هي عبارة عن مجموعة من مجموعة البيانات الضخمة جداً والمعقدة لدرجة أنه يُصبح من الصعب معالجتها باستخدام أداة واحدة فقط من أدوات إدارة قواعد البيانات أو باستخدام تطبيقات معالجة البيانات التقليدية " تعريف موسوعة الويكيبيديا: "

* عرف معهد ماكنزي العالمي سنة 2011 البيانات الضخمة أنها أي مجموعة من البيانات التي هي بحجم يفوق قدرة أدوات قواعد البيانات التقليدية من إلتقاط، تخزين، إدارة و تحليل تلك البيانات.

* ويعرفها (Grobelnik2012) بأنها" تشبه البيانات الصغيرة ولكنها أكبر وهذا يتطلب إجراءات مختلفة تتمثل في التقنيات والأدوات والبنية الهيكلية لتساعد في حل مشكلة البيانات الضخمة"

* وفي تعريف واتسون (Watson 2014) أنها مجموعة متنوعة من البيانات بكميات كبيرة يصعب التعامل معها بشكل جيد.

* ويعرف (Bieraugel 2016)[5] البيانات الضخمة بشكل عام بأنها البيانات التي لا يمكن تخزينها أو تحليلها من قبل الأجهزة والبرمجيات التقليدية.

* ويشير هلبرت (Hilbert 2013) إلى أن الفرص كبيرة عندما تقوم الدول والمنظمات بتحليل البيانات الضخمة إذ أن ذلك يساعدها في القدرة على اتخاذ القرارات في مجالات حيوية وهامة للمجتمعات خاصة ما يتعلق منها بالرعاية الصحية والتوظيف وزيادة الانتاج والحد من الجريمة ورفع مستوى الأمن.

[5] البيانات الضخمة واتخاذ القرار في جامعة الملك سعود: دراسة تقييمية لنظام اتقان -المؤلفون: د.علي بن ذيب الأكلبي1 - تاريخ النشر الكترونيا: 10 يناير 2019

2- خصائص البيانات الضخمة

تتميز البيانات الضخمة [6]بعدة خصائص يمكن اجمالها في الآتي:

• **الحجم(Volume)** : هو حجم البيانات المستخرجة من مصدر ما، وهو ما يحدد قيمة وإمكانيات البيانات لكي تُصنف من ضمن البيانات الضخمة، وبحلول العام 2020 سيحتوى الفضاء الإلكتروني على ما يقرب من 40.000 ميتابايت من البيانات الجاهزة للتحليل واستخلاص المعلومات.

• **التنوع(Variety)** : يُقصد به تنوع البيانات المستخرجة، والتي تُساعد المستخدمين سواء كانوا باحثين أو مُحللين على اختيار البيانات المناسبة لمجال بحثهم وتتضمن بيانات مُهيكلة(Structured Data) في قواعد بيانات وبيانات غير مهيكلة(Un Structured Data) مثل: الصور ومقاطع وتسجيلات الصوت وأشرطة الفيديو والرسائل القصيرة وسجلات المكالمات وبيانات الخرائط (GPS)، وتتطلب وقتاً وجهداً لتهيئتها في شكل مناسب للتجهيز والتحليل.

السرعة (Velocity) : يُقصد بها سرعة إنتاج واستخراج البيانات إرسالها لتغطية الطلب عليها حيث تعتبر السرعة عنصراً حاسماً في اتخاذ القرار بناء على هذه البيانات، وهو الوقت الذي نستغرقه من لحظة وصول هذه البيانات إلى لحظة الخروج بالقرار بناء عليها.• القيمة المضافة لهذه البيانات عندما يتم تحليلها ومعالجتها .

– **الموثوقية(Veracity)**: ويعنى: الدقة، والجودة والموثوقية للبيانات لأصحاب القرار، لذلك فإن البيانات الضخمة تتطلب أشكالًا وطرقًا جديدة لمعالجة بياناتها بما يعزز صناعة القرار الصحيح في المؤسسات المختلفة[7].

3- مصادر البيانات الضخمة:

• هناك مصادر تولدت من مدخلات أحد الأنظمة الآلية للمنظمة، كالسجلات الطبية الإلكترونية وزيارات المستشفيات وسجلات التأمين والسجلات المصرفية.

---

[6] مقال في منتدى أسبار الدولي حول " البيانات الضخمة ومجالات تطبيقها Featured " Monday, 03 December 2018 16:32

الدكتور/ عدنان مصطفى البار - أستاذ نظم المعلومات المشارك، قسم نظم المعلومات - جامعة الملك عبدالعزيز و الدكتور/ خالد علي المرحبي

أستاذ علوم الحاسب المساعد، قسم علوم الحاسبات -جامعة أم القرى

[7] مقال : أ.د. أحمد صادق عبد المجيد- أستاذ تكنولوجيا التعليم بكلية التربية جامعة سوهاج

• بيانات ناشئة عن التعامل بين جهتين مثل معاملات البطاقات الائتمانية والمعاملات التي تجرى عن طريق الإنترنت بوسائل منها الأجهزة المحمولة.

• بيانات تم التقاطها عبر شبكات أجهزة الاستشعار، على سبيل المثال، التصوير بالأقمار الصناعية، وأجهزة استشعار الطرق، وأجهزة استشعار المناخ ونحوها.

• مصادر واردة عبر أجهزة التتبع، مثل تتبع البيانات الصادرة عن الهواتف المحمولة والنظام العالمي لتحديد المواقع.

• تتبع البيانات السلوكية، مثل عدد مرات البحث على الإنترنت عن منتج أو خدمة ما أو أي نوع آخر من المعلومات، ومرات مشاهدة مقطع معين على الإنترنت.

• بيانات استطلاعات الآراء، مثل، التعليقات على وسائط التواصل الاجتماعي.

• البيانات الضخمة تزداد ضخامة بسرعة هائلة وتحتاج إلى خطة معالجة استراتيجية على مستوى العام .

• الفائدة من البيانات الضخمة لا زلات محدودة مقارنة بما فيها من فرص ونتائج يمكن الحصول عليها بعد التحليل.

الشكل رقم(1) مصادر البيانات الضخمة

4- مزايا استخدام أداة تحليل البيانات

- **السرعة** : القدرة على معالجة البيانات الضخمة بشكل أسرع
- **استهداف العميل** : القدرة على استهداف مجموعة محددة من الناس
- **الاحتياجات الاستباقية والتنبؤ:** القدرة على التنبؤ

- **القدرة على تحديد الفجوات:** القدرة على تحديد الفجوات في البيانات
- **تحديد المخاطر:** تحسين القدرة على تحديد المخاطر وتخفيفها
- **تحسين اتخاذ القرار:** بناءً على الفوائد المذكورة أعلاه، يمكن اتخاذ قرارات اكثر دقة

5- عيوب استخدام أداة تحليل البيانات

- **سوء جودة بيانات :** قاعدة (GIGO): إذا كانت البيانات التي تدخلها غير دقيقة
- **الحصول على البيانات :** القدرة على الحصول على البيانات خاصة من المصادر النوعية.
- **احتياجات التكلفة والأجهزة :** نفقات نقدية كبيرة، قد تحتاج البنية التحتية لتكنولوجيا المعلومات إلى التحديث
- **الأمان :**تحديث مستمر بسبب خطر الاختراق والتخريب.
- **التخزين :** يتطلب الحجم الكبير من البيانات سعة تخزين كبيرة
- **الاستعلام :** قد يجد المستخدم صعوبة في الاستعلام عن قاعدة البيانات.

6- الركائز الأساسية لتحليلات البيانات

- **الإطار المؤسسي**
  - التشريع
  - الحصول على البيانات الحساسة
  - اللائحة العامة لحماية البيانات
- **الإعداد التنظيمي**
  - من سيستخدم تحليلات البيانات، pa أو fa أو ca.
- **الكفاءات المهنية**
  - الموظفون الجدد
  - تدريب الموظفين الحاليين
- **الأدوات والتكنولوجيا**
  - سعر التكلفة المبدئي
  - الأجهزة المطلوب الحصول عليها
  - الترخيص

8- فوائد تحليل البيانات الضخمة:

- **زيادة التنافس بين المؤسسات والشركات** في مختلف مجالات عملها للعمل بشكلٍ أسرع والكفاءة في التحليل، الأمر الذي أدى الاعتماد على نظم تحليلية خاصة على أجهزة حاسوب تمتلك قواعد بيانات كبيرة، وقدرة عالية على الحوسبة والمعالجة.
- **تقليل التكاليف:** تخفف تقنيات تحليل البيانات الضخمة الحديثة من تكاليف تخزين الكميات الكبيرة من البيانات مثل مجموعة تخزين Hadoop الشهيرة، وتنظمها بشكلٍ مساعدٍ وأكثر فعاليةً للقيام بأعمال الشركات بشكلٍ سهلٍ.
- **اتخاذ القرارات بشكل أسرع:** نتيجة التحليلات السريعة والمفصلة لمصادر البيانات المُختلفة، تتمكن الشركات من فهم المعلومات وتحليلها بشكلٍ سريعٍ وفوريٍّ لاتخاذ القرارات المناسبة، والتي تحسن من آلية عملها وربحها في الأيام القادمة.
- **تقديم الخدمات والمنتجات الجديدة:** تتمكن الشركات من فهم وتحديد احتياجات العملاء ومعرفة ما يرضيهم، نظرًا لتحليل بيانات المنتجات والمبيعات وآراء الناس التي تقوم بها تقنيات تحليل البيانات، مما يجعل الشركات تقوم بتطوير المنُتجات الناجحة وتقديم المُنتجات والخدمات الجديدة التي ستنال إعجاب ورضى الزبائن.

9- التحديات التي تواجه المنظمات التي لديها بيانات ضخمة:

هناك العديد من التحديات والصعوبات التي تعاني منها المنظمات وهي تعمل على التعامل مع البيانات الضخمة، ويمكن إجمال أبرز هذه التحديات فيما يأتي:

- **حجم البيانات الضخمة المتزايد بصفة مستمرة.**
- **النمو الهائل والمتسارع في كمية البيانات.**
- **البحث والاسترجاع العشوائي داخل البيانات الضخمة.**
- **تنوع البيانات**
- **توفر الموظفين المتخصصين في تحليل البيانات الضخمة.**

• توفر الأنظمة الآلية الخبيرة التي تناسب احتياج المنظمة وتتمتع بقدرات جيدة ومرونة في الاستخدام والتطوير.

الشكل رقم(2) التحديات التي تواجه المنظمات التي لديها بيانات ضخمة

وبالتالي هنالك مجموعة من التحديات استخدام البيانات الضخمة[8] يمكن تلخيصها بما يلي :

1- **الإطار المؤسساتي:** وهي تلك المؤسسات المطلوبة لحماية أعمدة الديمقراطية كالسرية، وغالبا لا تطبق عندما يتعلق الأمر بالبيانات الضخمة، إذ يعتبر تحديا أساسيا يجب معالجته للرفع من استخدام البيانات الضخمة في التنمية المستدامة. والخصوصية هي حق الأشخاص في التحكم في المعلومات المرتبطة بها، والتي يمكن الكشف عنها داخل أسوار الديمقراطية. والحماية يجب تطبيقها لتفادي التعدي على هذا الحق في العصر الرقمي، إذ تعتبر الخصوصية تحدي شامل لأي شخص يريد استغلال البيانات الضخمة في التنمية، وذلك منذ أن أصبحت ممكنة التطبيق في جميع مجالات العمل بداية من اقتناء هذه البيانات وتخزينها إلى حفظها واستعمالها. في بعض الحالات، يرفع إنتاج البيانات نفسها من التحديات، خاصة مع لاوعي الأشخاص بحجم وأنواع البيانات المولدة يوميا. يجب إذن أن يُوفر إطار قانوني وقواعد أخلاقية لمشاركة البيانات التي يتم الموافقة على جمعها واستعمالها بدون وعي، وبدون فهم من أصحابها. لذلك، تعتبر هذه الخطوة مركز المجهودات المطلوبة للرفع من قدرة البيانات الضخمة في التنمية.

[8] .2 Maaroof, A. (2015). Big Data and the 2030agenda for sustainable development. the Development Goals in the Asia and the Pacific region, (pp. 52-1). Bangkok.

2019 ©Author(s), licensee HBKU Press.

2- الفجوة الرقمية: بالرغم من أن ثورة البيانات تنتشر عبر العالم بطرق متعددة، وبسرعات متفاوتة، فالفجوة الرقمية تتناقص بشكل أكبر مما توقع الكثيرون لاختلاف وفرة وأنواع البيانات الضخمة من بلد إلى آخر. فالبلدان التي تتعرض لأكبر نسب من الاختراق للهواتف الذكية والانترنت تنتج أكبر عدد من البيانات المولدة من خلال معلومات المواطنين. وتختلف البيانات المولدة حسب الفئات العمرية وحسب المداخيل الاقتصادية وحتى الجنس والموقع الجغرافي، وهذه الاختلافات والانحرافات يجب استخدامها في بناء السياسات القائمة على البيانات الضخمة. كما يجب أن يوجه اهتمام خاص للبلدان التي تنتج أقل نسب من البيانات الضخمة، أو تلك التي تمتلك قدرات ضعيفة في تحليل البيانات من أجل تفادي إضافة مؤثرات أخرى للفجوة الرقمية.

3- الوصول إلى البيانات الضخمة والشراكة: بالرغم من أن الكثير من البيانات المتاحة للعامة لديها استعمالات قوية في أهداف التنمية، إلا أن القطاع الخاص يحتفظ بأحجام كبيرة منها، حيث أن الكثير من الشركات تمانع مشاركة البيانات خوفا من المنافسة وحماية لزبائنها. لهذا، يتطلب العمل على البيانات الضخمة الشراكة بين منتجي البيانات ومستخدميها، وبين رؤساء تخزين وحفظ البيانات والمؤسسات لضمان تحقيق قوة البيانات الضخمة.

4- **التحديات التحليلية وتحديات القدرات:** إن عملية جعل البيانات الضخمة ذات معنى (لاستخراج المعلومات المناسبة) يتضمن مجموعة من المخاطر التحليلية التي يمكن أن تحد من صحة النتائج، حيث أن تحليل هذه البيانات لإعداد السياسات والتقييم يطرح الكثير من التحديات إما منهجية أو مرتبطة بصحة ترجمتها، أو المنهج المستخدم للتحليل من جهة، وتحدي القدرة على استعمال كل القوة التي تجلبها البيانات الضخمة بشكل فعال والذي لا يزال في خطواته الأولى بالإضافة إلى الإطار المؤسساتي الذي يزال يعيق تقوية قدرات العاملين به على استخدام وتحليل البيانات من جهة أخرى.

الجهاز المركزي للرقابة المالية

البحث الثالث : تقنيات تحليل البيانات الضخمة

1- تصنيف البيانات الضخمة (Big Data)

بشكل عام البيانات هي مجموعة من الحروف أو الكلمات أو الأرقام أو الرموز أو الصور (الخام) المتعلقة بموضوع معين. مثال على ذلك: بيانات الطالب (الأسماء – الأرقام المستوى – الشعبة – التخصص) بدون ترتيب، وينتج عن هذه البيانات بعد المعالجة ما يطلق عليه مصطلح معلومات.

و نميز نوعين من البيانات النوعية، و الكمية:

البيانات النوعية Qualitative Data تقيس صفة ما لظاهرة معينة دون أن تأخذ قيما عددية، وتنقسم إلى نوعين هما البيانات الإسمية و البيانات الترتيبية.

- **البيانات الإسمية Nominal Data**: تعتمد علي التصنيف النوعي للوحدات بغض النظر عن أهمية الترتيب. مثل الجنس يتكون من طبقتين ذكور و والإناث.
- **البيانات الترتيبية Ordinal Data**: تعتمد علي التصنيف والترتيب معا، لذلك يلعب الترتيب دورا أساسيا في تحديد معالم الظاهرة. مثل المستوى التعليمي(ابتدائي ، إعدادي، ثانوي، جامعي فأكثر)

البيانات الرقمية أو الكمية Quantitative Data تأخذ قيما عددية صحيحة أو كسرية حسب ظروف الحالة المدروسة. تنقسم أيضا إلى نوعين هما البيانات الدفترية والبيانات النسبية.

* **البيانات الدفترية Interval Data**:يتميز هذا النوع من البيانات بتساوي المسافات بين الرتب، ويستخدم هذا البيانات كثيرا في العلوم التربوية والنفسية والاجتماعية مثل قياس الذكاء وغيرها ، والجدير بالذكر أنه في هذا النوع من البيانات الصفر غير حقيقي، أي لا يعني الصفر في هذا النوع عدم وجود الخاصية فدرجة طالب تساوي صفر مثلا لا يعني أنه لا يعرف شيء في المقرر.

***البيانات النسبية Rational Data** : و تعني خاصية النسبية أن للصفر خاصية العدم أي خاصية انعدام الظاهرة مثل سرعة سيارة تساوي صفر تعني أن السيارة متوقفة.

2- آلية عمل تحليل البيانات

– في أغلب الأحيان، تستخدم الشركات منصات التخزين كمحطةٍ أوليةٍ لتخزين البيانات الضخمة قبل نقلها لقاعدة البيانات التحليليّة، وبمجرد أن تصبح البيانات كاملةً وجاهزةً، يتم تحليلها بوساطة برامج تحليل

عالية الجودة، والتي تعتبر كثيرةً في الأسواق، حيث تمتلك هذه البرامج أدواتٍ خاصةً للقيام بعملية التحليل، وأبرز هذه الأدوات:

- **التنقيب عن البيانات:** تقوم هذه الأدوات بتنقيح البيانات، بالإضافة إلى البحث عن جميع أصناف وأنماط البيانات.
- **التحليل التنبؤي:** تعتمد هذه الأدوات على الذكاء الاصطناعي في بناء نماذج الاحتمال والتنبؤ بالتطورات المستقبليّة، وبسلوك وحركة العملاء المتعلقة بآرائهم ومتطلباتهم.
- **التعلم الآلي:** يعمل على تحليل كمياتٍ كبيرةٍ من البيانات بالاعتماد على الخوارزميات المتقدمة.
- كما يلعب البرنامج المستخدم في استخراج النصوص وبرامج التحليل الرياضي والإحصائي دورًا كبيرًا في عملية تحليل البيانات الضخمة ككلٍّ، حيث يتم تصميم هذه البرامج بلغات البرمجة القوية والأساسية في مجال تحليل البيانات وبناء قواعد البيانات؛ مثل لغة بايثون Python، أو لغة R، وغيرها من اللغات، مثل لغة SQL، ولغة Scala، ويتم دعم هذه اللغات بوساطة تقنيات .SQL-on-Hadoop.

3- أنواع تحليل البيانات

1- **تحليل وصفي:** يُراد منه وصف ملخص للبيانات ولا يتطلب إيجاد تفسيرات لها، مثل ما يقدمه تحليل البيانات لإحصاء السكاني لبلد معين، حيث لا يقدم التحليل أكثر من خلاصة لما يشمله استبيان الاحصاء من جنس، وعمر، وعنوان وغيرها.

2- **تحليل استكشافي:** تحليل البيانات الاستكشافي يحاول إيجاد علاقات، اكتشافات، ارتباطات، ميول من القياسات لعدة متغيرات بغرض إيجاد أفكار وفرضيات معينة. مثال على التحليل الاستكشافي هو ما قام به مجموعة من الهواة الذين حللوا بيانات فضائية كثيرة جمعها مقراب كبلر فوجدوا نظاماً شمسياً من أربعة كواكب من خلال تحليل خصائص الضوء.

3- **تحليل إستنتاجي:** أحد أكثر تحليلات البيانات شيوعاً في البحوث العلمية، ويذهب إلى ما وراء التحليل الاستكشافي ليرى إن كانت الأنماط المكتشفة صالحة لكي تكون وراء مجاميع البيانات المتوفرة. مثال عليه كشف العلاقة بين التلوث البيئي ومتوسط العمر على مستوى الولايات في الولايات المتحدة. يقوم هذا التحليل بتقييس واحتساب العلاقات المختلفة بين القياسات المتوفرة.

4- **تحليل تنبؤي:** بينما يقوم النوع السابق بتقييس العلاقات واحتساب قيمها، يقوم التحليل التنبؤي بتوقع قياسات معينة من قياسات موجودة. مثلاً ما تقوم به مؤسسات الاحصاء في تنبؤ نتيجة الانتخابات من خلال تحليل سلوك التنبؤ الذي تتم ملاحظته في الاستبيانات.

5- **تحليل سببي:** يقوم هذا التحليل باحتساب مقاييس معينة في حال تغير مقاييس أخرى، مثلاً احتساب تأثير ممارسة طبية معينة على تقليل الإصابة بمرض معين.

6- **تحليل ميكانيكي:** يقوم التحليل السببي السابق بإيجاد علاقة لها نسبة معينة من الحدوث وعلى أثر بيانات قد تكون ضخمة جداً، مثلاً على مدى عقود تقول البيانات أن التدخين يؤدي إلى الإصابة بالسرطان، لكن الأمر ليس مؤكداً فقد لا تموت بالسرطان رغم تدخينك. ما يقوم به التحليل الميكانيكي هو إيجاد علاقة مؤكدة وحتمين بين قياسين.

ويهدف تحليل البيانات إلى اعداد ما يسمى بنموذج بيانات النظام. وتعتبر هذه العملية من الأنشطة الرئيسية لمرحلة التحليل وتتم نمذجة البيانات غالبا باستخدام النماذج البيانية ،أي المخططات والرسوم التي تشبه إلى حد ما مخططات تدفق البيانات

4- عملية تحليل البيانات

1- **تحديد متطلبات البيانات:** هي الخطوة الأولي في تحليل البيانات ويقصد بها التعريف والتحديد بنوعية وكمية وغيرها من الأشياء المهمة المطلوب توافراها في البيانات المراد تحليلها

مثال: البيانات المطلوبة هل هي ارقام، نصوص ام صور، هل ستكون البيانات محسوبة لشخص واحد ام لكل الاشخاص في هذا المكان، .... الخ من المتطلبات.

2- **تجميع البيانات:** يتم فيها تجميع البيانات من مصادر مختلفة بحيث تحقق المتطلبات في الخطوة الاولي . ومن الممكن ان يقوم بجمعها اشخاص، او الحصول عليها من خلال التقنيات الحديثة مثل الاقمار الصناعية، اشارات المرور، الانترنت ... الخ.

3- **تنظيم البيانات:** بعد مرحلة تجميع البيانات تبدأ عملية توزيع البيانات في شكل جداول لها صفوف واعمد كما في ملفات Excel.

4- **فحص البيانات:** من الضروري فحص البيانات حتي لا تكون المعلومات الناتجة بها اخطاء وغير صحيحة. ويتم ذلك من خلال مراجعة البيانات وازالة او تصحيح المغلوطة . البيانات المغلوطة قد تكون ارقام غير صحيحة ،بيانات مكررة ،بيانات مرتبات ولكن يوجد بها حروف ابجدية. ومن

الممكن التخلص من البيانات المغلوطة بإزالة المكرر واعادة حساب الارقام وفي عملية تدخيل البيانات نتأكد ان البيانات المدخلة لها نفس النوع لنفس العمود.

5- **اعداد النموذج المفاهيمي للبيانات:** تسمى هذه الخطوة أيضا نمذجة بيانات النظام. ويتم خلالها بناء النموذج الذي يعكس الموضوعات(الأشياء) الرئيسية للبيانات ، وعلاقاتها مع بعضها البعض .ويسمى التحليل في هذا المستوى بتحليل المضمون أو المعنى.

6- **تحليل العلاقات:** ويتم فيه تحسين النموذج المفاهيمي بإعادة تصميم الكينونات بطريقة تقلل التكرارات وتحول الكينونات إلى علاقات مبسطة يمكن التعامل معها بمرونة وسهولة .وتسمى هذه العملية أيضا تسوية أو تطبيع البيانات وبناء النموذج العلاقي للبيانات.

7- **تصميم قاعدة البيانات:** وتهتم بتحويل النموذج العلاقاني إلى توصيف قاعدة بيانات النظام .

الشكل رقم(3) عملية تحليل البيانات

5- الأطراف في منظومة البيانات الضخمة

لكى يتم تنظيم أي خدمة يجب تحديد الأطراف التي تتعامل مع هذه الخدمة وتحديد واجبات وحقوق كل طرف. تتكون منظومة البيانات الضخمة من عدة جهات تتفاعل في ما بينها ،هذه المنظومة مُكونة من موفر البيانات الضخمة ومُقدم خدمة البيانات الضخمة وعميل خدمة البيانات الضخمة ويمكن توضيح هذه الجهات كما يلي:

أولاً: موفر البيانات الضخمة : يعمل موفر البيانات الضخمة على توفير البيانات من مصادر مُختلفة إلى مُقدم الخدمة، وتشمل أنشطة موفري البيانات على سبيل المثال إنشاء البيانات وإنشاء المعلومات الوصفية(Meta data) التي تصف مصدر البيانات وإيجاد مصادر البيانات المفتوحة (open data) على الإنترنت وتوفير بيان الخدمات (Service catalogue) إلى مُقدم الخدمة عن البيانات القابلة للاستخدام.

ثانياً: مُقدم خدمة البيانات الضخمة : يقوم مُقدم الخدمة بتحليل البيانات الضخمة وتوفير البنية التحتية اللازمة لها وتشمل أنشطة مُقدم الخدمة على سبيل المثال البحث في مصادر البيانات وجمع البيانات عن طريق الطلب المباشر من موفر البيانات أو البحث في الإنترنت (Data Crawling) كما تشمل أنشطة مقدم الخدمة تخزين البيانات ودمجها وتوفير الأدوات لتحليلها ودعم إداراتها مثل خصوصية البيانات وأمن البيانات وملكية البيانات وغيرها.

ثالثاً: عميل خدمة البيانات الضخمة : وهو المستخدم النهائي لمنظومة البيانات الضخمة أو هو نظام يستخدم النتائج أو الخدمات التي يُقدمها مُقدم خدمة البيانات الضخمة، كما يمكن للعميل أن يُنتج خدمات جديدة أو معرفة وذلك إعتماداً على نتائج تحليل البيانات الضخمة. تشمل أنشطة العميل على سبيل المثال طلب خدمة البيانات الضخمة من مزود الخدمة واستخدام مُخرجات خدمة البيانات الضخمة في النشاط الذي يرغبه ويتخصص فيه.

البحث الرابع : تطبيقات تحليل البيانات الضخمة(Big Data)

1-تحليل البيانات الضخمة(Big Data) وتحسين التعليم

أدى استخدام أدوات التعلم عبر الإنترنت والبرامج القائمة على التفاعل بصورة متزايدة في مجال التعليم إلى زيادة حجم البيانات، واختلاف نوعية البيانات الكبيرة التي يُمكن جمعها من بيئات التعلم، فهنا نجد بيانات كبيرة عن المتعلمين، وخبرات التعلم لدى المتعلمين، كما نجد بيانات متعمقة داخل بيئات التعلم، وبيانات حول التفاعلات الاجتماعية في بيئات التعلم، وبيانات مُفصلة عن أنشطة التعلم من نصوص ووسائط ومقاطع فيديو وغيرها، كما تختلف هذه البيانات في نوعيتها وعمقها بنسب متفاوتة.

يمكن الاستفادة من تحليل هذه الأنواع من البيانات الضخمة في التعليم، لتوفير مجموعة مُتنوعة من الفرص والخيارات بهدف تحسين تعلم الطلاب من خلال التعلم التكيفي أو التعليم القائم على الكفاءة، مما ينتج عنه تعلم أفضل نتيجة لتشخيص أسرع وأكثر تعمقاً في بيانات حقيقية تراكمية لاحتياجات التعلم أو المتاعب التي تواجهه أثناء عملية التعلم، بما في ذلك تقييم المهارات مثل التفكير المنظم، والتعاون، وحل المشاكل في سياق عميق، وتقييم 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 أدوات حديثة وفعالة لقياس أداء الطلاب للمهام التعليمية، ويمكن أن تُساعد كذلك في تصميم بيئات تعلم تصميماً مُخصصاً وفق احتياجات مُحددة للطلاب، ويمكن أن تعطي تحليلاً واضحاً لردود الفعل الفردية والجماعية لمجموعة من القضايا التعليمية وغيرها من المميزات.

2-تحليل البيانات الضخمة لتحسين عملية صنع القرار

إن عملية اتخاذ القرارات تُعد محور العملية الإدارية وجوهرها وإن نجاح المؤسسة أو القطاع الحكومي يتوقف إلى حد كبير على قدرة وكفاءة القيادة الإدارية على إتخاذ القرارات الإدارية المناسبة. إن عملية صنع القرار تبدأ بتجميع البيانات ومُعالجتها واستخلاص المعلومات التي بناء عليها يتم اتخاذ القرار حيث بدأت تعتمد العديد من الشركات الكبيرة والقطاعات الحكومية على سياسة تحليل البيانات الضخمة والمعقدة والتي تحتاج إلى البرمجيات المتخصصة في مجال إدارة البيانات والتحليلات، والتي لا يمكن مُعالجتها باستخدام أداة واحدة فقط أو العمل على تطبيقات مُعالجة بيانات تقليدية، فمن المعروف أن جمع البيانات والمعلومات تُساعد على التوصيف الدقيق للمشكلة وتحليلها للوصول إلى نتائج دقيقة، لذلك كان لابد من اعتماد نظام إداري يشمل تحليل البيانات الضخمة والهائلة جداً. يستخدم القطاع الحكومي والشركات الكبيرة نظام تحليل البيانات الضخمة لتحسين العمليات الداخلية، مثل إدارة المخاطر، إدارة علاقات العملاء، والخدمات اللوجستية. كما يستخدم لتحسين المنتجات والخدمات القائمة، وتطوير الخدمات والمنتجات الجديدة والاستفادة من المعلومات وتقديم العروض المناسبة للعملاء في الوقت المناسب.

كيف تساعد البيانات الضخمة المؤسسات على اتخاذ قرار أفضل؟
تحويل 12 تيرابايت من التغريدات يوميا إلى تحليل لمشاعر الناس حول منتج ما
تحويل 350 مليار متر من القراءات السنوية إلى تحليل لتوقعات استهلاك الطاقة
تحليل تفاصيل سجلات 500 مليون مكالمة يومية في الوقت الحقيقي للتنبؤ بتوجه العملاء بشكل أسرع
التدقيق على 5 ملايين معاملة تجارية يوميا لكشف أي عملية احتيال محتملة
مراقبة البث المباشر لـ 100 كاميرا مراقبة لتحديد النقاط التي تجذب الاهتمام

توصلت الدراسة البيانات الضخمة واتخاذ القرار د. علي بن ذيب الأكلبي[9] في جامعة الملك سعود إلى عدة نتائج وتوصيات يمكن إجمالها كما يأتي:

- البيانات الضخمة تزداد ضخامة بسرعة هائلة وتحتاج إلى خطة معالجة استراتيجية
- الفائدة من البيانات الضخمة في الجامعة لا زلات محدودة مقارنة بما فيها من فرص ونتائج يمكن · الحصول عليها بعد التحليل.
- أظهرت الدراسة القيمة التجارية الكبيرة للبيانات الضخمة وأن عوائدها متنوعة بناء على آليات · وطرق تحليلها.
- أظهرت الدراسة أن خصوصية البيانات الخاضعة للتحليل وقضايا الملكية الفكرية للبيانات هو

– أظهرت الدراسة أن تحليل البيانات الضخمة يساعد في الكشف عن الاتجاهات الاجرامية أو التجارية ·أو السلوكية أو الوظيفية لدى اصحاب هذه البيانات مما يجعل تحويلها إلى منتج مفيد للغاية ضمن دائرة الاقتصاد المبني على المعرفة أمرا واقعا.

– أظهرت الدراسة أن تقنيات الحوسبة السحابية من أهم الحلول والطرق المفيدة في ايجاد · مساحات تخزينية للبيانات الضخمة وتساعد في التعامل معها بالسرعة المناسبة وأوصت الدراسة بضرورة سن الانظمة والتشريعات المنظمة على المستوى الوطني والإقليمي والدولي بشكل واضح لموضوع الملكية الفكرية وخصوصية البيانات.

[9] البيانات الضخمة واتخاذ القرار في جامعة الملك سعود: دراسة تقييمية لنظام اتقان , Journal of Information Studies and Technology 2018:2,15 https://doi.org/10.5339/jist.2018.15

3-الإستفادة من البيانات الضخمة في المجال الاقتصادي :

لقد أصبح بإمكان الشركات والمؤسسات والهيئات اليوم على إختلاف أنواعها تحليل حركة العملاء من شراء وبيع ونحوه بدقة أكبر ليتمكنوا وفقاً لذلك من معرفة السلع الأكثر طلباً أو تلك الراكدة ويقترحوا على عملائهم سلع معينة وفقاً لعمليات الشراء التي تتم. كما أصبح لديهم القدرة على فهم سلوك العملاء بشكل أكثر دقة وتحديد المميزين منهم ومن هم بحاجة لمساعدة أو لتحديد توجهاتهم أو مُراقبة أدائهم. هذا الأمر ليس فقط لمراكز البيع التقليدية بل يشمل المتاجر الإلكترونية على شبكة الإنترنت وعلى نطاق أوسع.

فأصبح يتفاجأ مُستخدم شبكات التواصل الإجتماعي أو البريد الإلكتروني في أحيان كثيرة، بظهور إعلانات تجارية لسلع قام مسبقاً بالبحث عنها في تطبيقات أخرى، بل أكثر من ذلك هناك بعض الخوارزميات التي تستخدم بيانات تحديد الموقع على جهاز الهاتف لاقتراح الإعلانات. ومن هنا نجد أن ذلك يحدث نتيجة لتحليل البيانات الضخمة الناتجة من هذه المواقع والإستفادة منها في التسويق وذلك باستخدام كل جزء صغير من البيانات المتاحة عن المستخدمين لمعرفة ميولهم وتفضيلاتهم بغية عرض البضائع بأمثل طريقة ممكنة تجلب لشركات التسوق الإلكتروني أعظم ربح ممكن. ولا تقوم بتحليل البيانات التي تحصل عليها من تصفح المستخدم للإنترنت وحسب، بل قد تتبع زيارته للأسواق الحقيقية من خلال جهاز تحديد المواقع المُثبت بجهازه. ربما يعتقد البعض أن هذا إختراق للخصوصية، لكن في حقيقة الأمر فإن الشركات التي تقوم بذلك قد حصنت نفسها من خلال اتفاقية الشروط التي يوقع عليها المستخدم عند تسجيل الدخول إلى تطبيقات التواصل الإجتماعي كالفيسبوك وتويتر.

وقد دخلت عملية التجارة الإلكترونية في السنوات الأخيرة مستوى جديداً من التنافس، بظل الكم الهائل من بيانات المستخدمين التي توفرها شبكات التواصل الإجتماعي واستخدام الإنترنت. أصبح الشاغل الأول لعمالقة التجارة الإلكترونية هو كيفية البقاء في المنافسة. فمن ناحية، تحتاج الشركات للترويج لبضائعها وهذا يتطلب معرفة باحتياجات الزبائن، ومن ناحية أخرى تحتاج لأن تقدم بضائعها بأسعار تنافسية في نفس الوقت تضمن لها هامش ربح كبير.

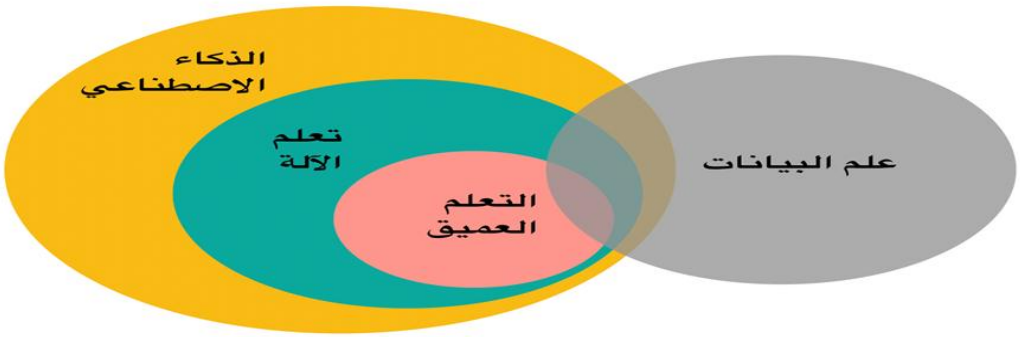
البحث الخامس : الذكاء الاصطناعي وعلم البيانات

يُعد مفهوم الذكاء الاصطناعي وعلم البيانات واحدًا من أبرز المفاهيم التي حققت انتشارًا كبيرًا منذ بداية العقد الثاني من القرن الحالي؛ حيث دخل هذا المفهوم في كل المجالات؛ ما جعله جزءًا أساسيًا من حياتنا اليومية؛ لكونه يمتلك الإمكانيات التي توفر كل احتياجاتنا؛ لذلك سنتناول في السطور التالية مفهوم الذكاء الاصطناعي وعلم البيانات[10].

تعود بداية ظهور مفهوم الذكاء الاصطناعي إلى عام 1950 عندما قدم العالم "Alan Turing" اختبار "تورينج" الذي كان يهدف إلى تقييم الذكاء لجهاز الكمبيوتر وتصنيفه ذكيًا في حال قدرته على محاكاة العقل البشري؛ حيث جاء هذا المفهوم حاملًا معه الكثير من الإنجازات الكبيرة التي تُسهل على البشر حياتهم، ويتميز الذكاء الاصطناعي بالقدرة الفائقة على التفكير وتحليل البيانات وحل المشكلات واتخاذ القرارات المناسبة والفعالة ببعض المجالات.

وبطبيعة الحال، تعتمد عملية التشغيل في أنظمة الذكاء الاصطناعي على خوارزميات باستخدام تقنيات التعلم الآلي والتعلم العميق؛ حيث تتولى تلك الخوارزميات تغذية بيانات أجهزة الكومبيوتر لأنظمة الذكاء الاصطناعي؛ وذلك من خلال تقنيات إحصائية لتمكين أنظمة الذكاء الاصطناعي من التعلم الآلي.

ورغم أن الذكاء الاصطناعي يقدم صورًا عن الروبوتات عالية الأداء الشبيهة بالإنسان، إلا أنه لا يهدف إلى أن يحل محل البشر على الإطلاق، ولكنه جاء لتعزيز القدرات والمساهمات البشرية بشكل كبير.



الشكل رقم(4) الذكاء الاصطناعي وعلم البيانات

[10] 11- اسلام النجار 11 أكتوبر، 2020 - - الذكاء الاصطناعي وعلم البيانات..

يمكن **تعريف الذكاء الاصطناعي** بأنه" عبارة عن مجموعة من الأنظمة تهدف إلى جعل الآلات والحواسيب الرقمية والتقنيات الحديثة قادرة على تحقيق أهداف معينة بطريقة مشابهة للبشر أو تتعدى قدرة البشر في أغلب الأحيان ". بمعنى آخر هو عبارة عن أنظمة تُحاكي الذكاء البشري لأداء المهام والتي لديها القدرة على أن تُحسن من نفسها باستخدام المعلومات التي تجمعها، فمثلًا قبل ظهور الذكاء الاصطناعي كانت برامج ألعاب الشطرنج تُصمم عن طريق تخزين أكبر عدد ممكن من الاحتمالات والنقلات التي يُمكن أن يقوم بها الاعب الذي يُمارس اللعبة من خلال جهاز الحاسوب، ولكن في ظل انتشار الذكاء الاصطناعي أصبحت الأمور أكثر تطورًا؛ حيث يتم تغذية برنامج لعبة الشطرنج بالخوارزميات المتخصصة التي تكون مهمتها تزويد القدرة على التعلم والاستنتاج وتقييم ردود الأفعال في كل مرة يقوم فيها اللاعب بممارسة اللعبة مع لاعب جديد.

تُكمن أهمية الذكاء الاصطناعي في حماية البشر من المخاطر والضغوطات النفسية ومساعدتهم في التركيز على الأشياء المهمة فقط؛ وذلك عن طريق الآلات التي تنفذ المهام الخطرة واستكشاف الأماكن المجهولة والمشاركة في عمليات الإنقاذ أثناء الكوارث الطبيعية.

وفقًا لتقارير أصدرتها شركة "إنتل" فإن حجم البيانات التي أنتجها العالم منذ بداية عصر الإنترنت وحتى عام 2006 يقدر بنحو أكثر من 8 إكسا بايت أي ما يُعادل 8 بلايين جيجا بايت، وتضاعف هذا الرقم أكثر من 1500 مرة خلال عام 2018، ليصل إلى 5.6 زيتا بايت "الزيتا بايت يعادل الألف بليون جيجا بايت"، ويتوقع بعض الخبراء أن يتضاعف هذا الرقم 5 مرات بنهاية عام 2022م.

وأدت كل هذه البيانات الضخمة إلى توجه الكثير من دول العالم نحو إدماج سياسة البيانات في حوكمتها من أجل بناء مجتمعات تعتمد في الأساس على المعرفة، وفي هذا الصدد قامت الولايات المتحدة الأمريكية بتأسيس بوابة إلكترونية تحت عنوان "داتا.جوف "data.gov تضم نحو 200 ألف قاعدة بيانات متوفّرة للعلن.

وتعتمد الحكومات على تقنيات الذكاء الاصطناعي التي لديها القدرة على التعلم والاستنتاج ورد الفعل لتحليل البيانات ومعالجتها بالشكل الصحيح ومساعدتها في اتخاذ القرارات الصحيحة، إذًا يُمكننا القول إن تقنيات الذكاء الاصطناعي لديها القدرة الفائقة على تحليل البيانات الضخمة والتنقيب عنها.

البحث السادس : نظم وبرامج الإلكترونية لتحليل البيانات

وفي ظل هذه الظروف التي أدت إلى تضخم البيانات وسرعة تواردها معتمدة على ما أسهمت به مستجدات التقنية وأدواتها وبيئات العمل الرقمية ووسائل التواصل الاجتماعية أصبحت الحاجة ملحة الى بناء أو اقتناء نظام يضمن سرعة فائقة في تحليل البيانات الضخمة في الوقت المتزامن، وهذا ما أدى إلى ظهور العديد من النظم التي اختصت بأعمال التحليل للبيانات[11] مثل:

- **برنامج spss** من البرامج الإحصائية الهامة التي يعتمد عليها الكثير من الباحثين ومن الجدير بالذكر إن النسخة الأولى من البرنامج تم إنشاءها في عام 1968 ويعمل ذلك البرنامج على التحليل الكمي. يعتمد برنامج spss على الجداول والبيانات الرقمية وتوزيع الإحصاء الوصفي بطريقة سهلة ومتاحة للجميع.
- **برنامج stata** يتسم برنامج stata برنامج يعتمد على البيانات التفاعلية التي تعتمد على منصات إلكترونية تقوم بجميع العمليات الإحصائية المعقدة ويتميز بواجهة تحكم يتم تحريكها باستخدام زر الماوس. كما يمكنك إنشاء البيانات والجداول البيانية ومجموعة من الصور التوضيحية.
- **برنامج SAS** يعرف برنامج sas بالتحليل الإحصائي والذي يعمل على كتابة التقارير وإنشاء المخططات والقدرة على إدارة المشروعات والأشياء من خلاله.
- برنامج إنفو يعتمد برنامج إنفو على برمجيات دقيقة للعمل على انشاء المعلومات والتحليل والقدرة على بناء المعلومات والبيانات مهما كان حجمها. كما يتسم ذلك البرنامج بتحليل المهام بشكل يدوي والعمل على تصنيفها وفرزها واكتشاف جميع الاتجاهات والاختبارات.
- **برنامج Hadoop** وهو برنامج مفتوح المصدر و يعتبر من أفضل أنظمة إدارة وتحليل البيانات الضخمة ويتميز بإطار العمل الذي يسمح بتجهيز مجموعات كبيرة من البيانات موزعة عبر مجموعات من أجهزة الكمبيوتر التي تستخدم نماذج برمجة بسيطة. وهي مصممة لتتدرج من خوادم واحدة إلى الآلاف من الآلات، وكل آلة تقوم بالتخزين المحلى. فبدلا من الاعتماد على الأجهزة لتقديم خدمة توفر عالية، فالمكتبة نفسها مصممة لكشف أي فشل والتعامل معه في حيز التطبيقات، وذلك بتقديم خدمة عالية ومتاحة على رأس مجموعة من أجهزة الكمبيوتر، كل واحدة منها قد تكون عرضة للفشل. ويعتبر برنامج أو منصة برمجية مفتوحة المصدر مكتوبة بلغة الجافا تستخدم لتخزين ومُعالجة البيانات الضخمة بشكل موزع أي أن تخزين هذه البيانات الضخمة تكون على عدة أجهزة ومن ثم توزع عملية المعالجة على هذه الأجهزة لتسريع

[11] افضل 10 أدوات لتحليل البيانات - By A -WINDOWS 7Home 0la Home دليل التقنيةWindows

نتيجة المعالجة وتعود أو تستدعى كحزمة واحدة. تتكون الأدوات التي تتعامل مع البيانات الضخمة من ثلاثة أجزاء رئيسية وهي: أدوات التنقيب عن البيانات(Data Mining) وأدوات التحليل(Data Analysis) وأخيراً أدوات عرض النتائج(http://hadoop.apache.org). (Dashboard).

– **برنامج SAP HANA** يساعد هذا النظام مستخدميه في القيام بعمليات التحليل المتزامن للبيانات الضخمة والتي ترد إلى منصة النظام، وهذا يمكن المنظمات التجارية والغير تجارية التي تقتنيه من إنجاز العمليات التجارية بشكل أسرع من خلال ما توفر لها من بيانات معالجة تساعد في اتخاذ القرار والقيام بمهام التخطيط والتنفيذ بكفاءة عالية،(https://hana.sap.com/abouthana.html)

الشكل رقم(5) نظم وبرامج الإلكترونية لتحليل البيانات

وتم تصميم برامج تحليل البيانات لتسريع عملية تحويل البيانات الخام إلى معلومات مفيدة عالية الجودة في بضع نقرات. ومع ذلك ، فإن السوق تغمره أدوات تحليل البيانات مما يجعل من الصعب الاختيار من بينها. وبالتالي ، من المهم البحث عن عدد قليل من الأشياء في الأداة المناسبة مثل أدائها ومنحنى التعلم وتحديد التكلفة الإجمالية للملكية الخاصة بها ، إلخ.

excel data analytics : إنه يتكيّف أنماطك ووفقًا لذلك يرتب بياناتك لتوفير وقتك وجهودك.

يقدم حسابات وقوالب حديثة لتجربة إنشاء ورقة مخصصة. وتأتي محملة بأحدث الرسوم البيانية لمساعدتك في عرض بياناتك بطريقة جذابة. ويخلق التنبؤات في نقرة واحدة فقط للتنبؤ بالاتجاهات ، دون عناء.

خيار لمشاركة جدول البيانات الخاص بك والتعاون في الوقت الحقيقي. علاوة على ذلك ، يساعدك على إنشاء نماذج باستخدام ميزة تحليلات الأعمال المتقدمة.

Zoho Analytics : تتيح أداة تحليل الأعمال وتحليلات الأعمال ذاتية الخدمة ، Zoho Analytics ، إنشاء عروض بيانات جذابة ولوحات معلومات ذكية في لمح البصر. يساعدك على تحويل أجزاء كبيرة من البيانات الخام إلى معلومات قيمة في شكل تقارير وفي لوحات المعلومات.

- يتيح سحب البيانات من جداول البيانات والملفات المسطحة (الملفات النصية ، CSV، HTML، Excel ، إلخ) وتحميلها.
- يسهل التحليل المرئي للبيانات وإنشاء تقارير ولوحات معلومات متعمقة باستخدام واجهة السحب والإفلات.
- يوفر مجموعة من واجهات برمجة التطبيقات (API) للويب الديناميكية المدعومة من HTTP والتي تساعد في تكامل البيانات ومعالجة بيانات mete والمزيد
- يضمن أن بيانات آمنة مع دعم قوي للتكرار والسرية والخصوصية والنسخ الاحتياطي.

Tableau Public : عبارة عن منصة مجانية لكنها قوية لأولئك الذين يبحثون عن أداة مجانية لتحليل البيانات. يتيح لك تمثيل ونشر البيانات حول أي موضوع يهمك. يمكّنُك IT من تصميم الرسوم البيانية التفاعلية والخرائط. ويساعد Tableau Public على توصيل تنسيقات مختلفة تتضمن MS Excel و MA Access وأنواع مختلفة أخرى من الملفات النصية ومساحة تخزين البيانات محدودة بـ 10 جيجابايت.

وتقتصر صفوف البيانات على 1 ، 000،000 في ملف واحد كونه نظامًا أساسيًا مفتوح المصدر ، فإن جميع المحتويات المحفوظة على النظام الأساسي مفتوحة للجمهور ، عبر الإنترنت مع خيار تنزيل نسخة من مصنف البيانات من قبل أي شخص.

يتمتع بالمرونة وبدقائق يتم نشره في أي مكان على الويب ، تعد Tableau Public واحدة من أدوات التحليل الكاملة الأكثر ديناميكية وآمنة وقابلة للتكيف.

IBM Analytics : ينتمي IBM Analytics إلى واحد من القادة الموثوق بهم الذين يسرعون عملية تحويل البيانات الخاصة بك باستخدام الذكاء الاصطناعي والمقاربة الكلاسيكية. تغطي جميع أنواع الصناعات مع تمكينك من اتخاذ قرارات ذكية أثناء التحليلات.

- يبسط عملية جمع البيانات ، والترتيب ، والطريقة التي تقوم بتحليل البيانات من أجل تحسين اكتسابها وإدارتها وحجمها.
- حرية جمع جميع أنواع البيانات من أنواع مختلفة من مصادر البيانات.
- خيار لتوسيع نطاق معرفتك عن طريق تضمين رؤى في قرارك استنادًا إلى الدليل والذي لم يكن الوصول إليه ممكنًا في وقت سابق.
- استخدام التعلم الآلي الذي يسرع دمج مشاريع علوم البيانات من خلال تحسين الذكاء .
- يسمح بالتحليل التقليدي بمراجعة قيود العمل وتعظيم المفاضلات التجارية.
- يحصل التحليل التنبئي على التنقيب عن البيانات أو النمذجة التنبؤية أو قوة التحليلات المتقدمة أو تحليلات النص في مكان واحد.

Python : إنها لغة برمجة مفتوحة المصدر وعامة الغرض مخصصة لعالم الترميز. فهو يساهم في البرمجة المستندة إلى الكائنات ، والهياكل والبرمجة الوظيفية ، وأكثر من ذلك. في الآونة الأخيرة ، يقدم واجهة برمجة تطبيقات متقدمة لـ AI والتعلم الآلي.

- يحدد بوضوح الوظائف مع السماح لقوائم الحجج الإجبارية والبديلة ، وكذلك الكلمات الأساسية والحجج العشوائية.
- مدعوم بـ "قوائم" التي تعد واحدة من أنواع البيانات المركبة المستخدمة في بيثون ، والتي يمكن سردها ، تجزئة ، والتأثير مع وظائف متكاملة أخرى.
- يجعل العمليات الحسابية وصيغة التعبير بسيطة من خلال السماح للعوامل + و – و * و / ووظائفها بشكل طبيعي. للسماح باستخدام الأقواس () للتجميع.

Apache Spark : هذا البرنامج المفتوح المصدر مصمم خصيصًا للمطورين ، ويدعم معالجة SQL والدفق والدُفعات. إنه وسيلة سهلة لاستخدام أداة تحليل متكاملة مدمجة بسرعة ولديها القدرة على تحويل أجزاء كبيرة من البيانات.

- يتيح كتابة التطبيقات في Java و Python و SQL و Scala و R.
- مزود بمكتبة ضخمة تضم SQL و DataFrames و MLlib للتعلم الآلي و GraphX و Spark Streaming ، وكلها يمكن دمجها بسلاسة في التطبيق.
- يعمل مع Hadoop أو Apache Mesos أو Kubernetes أو مستقل أو سحابة ، وكذلك الوصول إلى أنواع مختلفة من مصادر البيانات.

R Programming : يعد R Programming أداة تحليل بيانات مجانية مثالية للمهام الإحصائية وتصور الرسومات وإعداد التقارير. يعتمد على لغة كمبيوتر مترجمة تسمح بالفروع والحلقات والبرمجة المعيارية مع الوظائف. يتكامل أيضًا بسلاسة مع الأكواد المكتوبة بلغات C أو C ++ أو .Net أو Python أو FORTRAN.

- لغة برمجة سهلة وفعالة تأتي محملة بحلقات وخيارات مخرجات وشروط مشروطة ووظائف متكررة يحددها المستخدم.
- يعالج البيانات بفعالية ويوفر مساحة تخزين.
- يوفر مجموعة واسعة من الأدوات لحساب القوائم والمتجهات والمصفوفات.
- يسهل مجموعة ضخمة ومنطقية وموحدة من الأدوات المصممة خصيصا لتحليل البيانات.
- يوفر إمكانات رسومية وخيار لتقديم إخراج تحليل البيانات مباشرة على شاشة الكمبيوتر أو على الأوراق المطبوعة.
- يعمل على أنظمة متعددة مثل Windows و Mac OS و Unix.

Splunk : يوفر بشكل أساسي عرضًا شاملاً للبيانات المتصلة من جميع الأنظمة والأجهزة والتفاعلات في مؤسستك. ويتيح Splunk الاطلاع على ما يحدث في الوقت الفعلي. يساعدك هذا على إنشاء تأثير في شركتك عبر الكشف عن القوة الحقيقية لبياناتك. علاوة على ذلك ، فإنه يتيح لك تقديم رؤى قابلة للتطبيق لأولئك المعنيين ، في الوقت الفعلي باستخدام الواقع المعزز للهاتف المحمول.

- يسمح لك باستجواب Splunk للوصول إلى بيانات المستخدم دون الحاجة إلى كتابة SPL وهو أمر مفيد حتى بالنسبة لأولئك الذين لديهم خلفية غير تقنية.
- خيار لمسح رموز QR أو علامات NFC مما يسمح لك باختبار بياناتك ولوحات بيانات Splunk تلقائيًا على الكائنات مع تقديم رؤى قيمة لـ Splunk للمستخدمين الذين لا ينتمون إلى SPL.
- توفر لوحات معلومات مصممة خصيصًا للأجهزة المحمولة ، وتتلقى وتتخذ إجراءات بشأن الإشعارات في نفس الوقت.
- يوفر خدمة سحابية يمكن الاعتماد عليها ومحمية بكلمة مرور تمكن الأجهزة المحمولة من التفاعل مع المثيلات المدمجة والمستندة إلى مجموعة النظراء.
- بشكل أساسي ، يسمح لك باتخاذ قرارات أسرع في الوقت الفعلي عبر العمل.

Qlik Sense : يعد Qlik Sense معيارًا لأحدث سلسلة من منصات تحليل البيانات. لأنه يأتي مع برنامج تحليلات ذات صلة فريدة من نوعها ، كاملة مع AI المكررة ، وبيئة متعددة السحابية التي يمكن تحجيمها. مع كل هذه الوظائف وأكثر من ذلك فإنه يتيح للأشخاص في مؤسستك اتخاذ قرارات أسرع وأفضل على أساس يومي.

- مزودة بمرفق السحب والإفلات المرئي ، ووظيفة البحث الذكي.
- اتصالات سريعة ويمكن الاعتماد عليها لأنواع مختلفة من مصادر البيانات.
- يتيح الوصول إلى تحليلات البيانات في الوقت الفعلي في أي وقت وفي أي مكان.
- ميزات مرفق سرد القصص وكذلك راحة الخدمة الذاتية سهلة.
- مزود بمكتبة كائنات مشتركة ، وواجهات برمجة التطبيقات الديناميكية المفتوحة والمعيارية.
- يوفر سهولة التطوير التدريجي مع مرفق توحيد البيانات متعدد المصادر الذي يتضمن بيانات كبيرة.

الكفاءات اللازمة لتحليل البيانات

المهارة الأساسية لتحليل البيانات هو القدرة على فهم معنى تلك البيانات، وعليه يرى فريق التدقيق على أهداف التنمية المستدامة [12]أنه من المفيد توظيف محلل ل لبيانات للعمل ضمن الفريق لتنفيذ عمليات التحليل

[12] نموذج مبادرة تنمية الانتوساي للتدقيق على أهداف التنمية المستدامة isam - النسخة الاسترشادية - 2020

وإنشاء ا لتصوا رت. قد ترغب الأجهزة العليا للرقابة في الاستثمار في الكفاءات التالية للحصول على أساس قوي لتحليل البيانات .

- **الحدس في البيانات**: القدرة على فهم البيانات المنتظمة وغير المنتظمة بشكل حدسي .
- **جمع البيانات والمعالجة المسبقة:** القدرة على تصميم عمليات الجمع أو آليات التجميع والمعالجة المسبقة مثل التنقية والتجميع والتصفية عند الضرورة.
- **التحليل الإحصائي وتفسير البيانات**: القدرة على اجراء عمليات التحليل فردية أو متعددة المتغيرات وتفسير نتائج التحليل للعثور على العلاقة الاعتمادية أو العلاقات (السبب – الأثر) أوالتوجهات .
- **عرض البيانات (التقارير والتصور):** القدرة على اختيار الطريقة الصحيحة للإبلاغ عن المعلومات وإعداد التصورات .
- **معرفة الأدوات:** excel, idea

كما قد تشمل بعض الكفاءات المتقدمة على ما يلي:

- **الحصول على البيانات وحفظها**: القدرة على جمع قواعد البيانات واستعادتها محليًا وإدارة إمكانيات الوصول إلى البيانات .
- **الاستعلام:** تصميم استعلامات قاعدة البيانات لاستخراج البيانات من قواعد البيانات المنظمة .
- **البرمجة:** تطوير الخوارزميات لحل متطلبات تحليل البيانات .
- **المعرفة بالأدوات** : tableau, r, python, postgresql
- بصرف النظر عن قائمة المهارات الأساسية، قد يرى الجهاز الأعلى للرقابة أنه من المفيد الاستثمار في التعلم الآلي ومها ا رت الذكاء الاصطناعي مع وضع استراتيجية طويلة المدى في الاعتبار.

الفصل الثاني : رقابة البيانات الضخمة (Big Data)

البحث الاول : الرقابة الداخلية والبيانات الضخمة

غالباً ما تؤدي البيانات والتحليلات الضخمة إلى تغيير جذري في طريقة عمل الشركات، ولا يمكن للمراجعة الداخلية أن تتخلف عن الركب، فخدمات المراجعة الداخلية التقليدية مستمرة بشكل جيد منذ عقود، ولكنها

تحتاج إلى تحديث / تنقيح لدمج التحليلات من أجل الاستفادة بنجاح من إمكاناتها، والواقع يظهر عددا متزايدا من الحالات التي أصبحت فيها التحليلات جزءا لا يتجزأ من جوهر منهجية المراجعة الداخلية، فدمج تحليلات البيانات في التدقيق الداخلي يمكن أن يسفر عن تحسينات كبيرة في السرعة والدقة، ولكنه يتطلب تغييرا جذريا في العقلية والنهج.

إن نشاط المراجعة الداخلية يجب ان يساعد الشركات في الحفاظ على ضوابط رقابية فعالة من خلال تقييم فعالية وكفاءة هذه الضوابط ومن خلال تعزيز التحسين المستمر، وبالتالي يجب ان تقوم المراجعة الداخلية بتقييم الضوابط الرقابية على العمليات والتكنولوجيا، كما يجب أن يركز المراجعين الداخليين بشكل كبير على كيفية استهلاك البيانات والإجراءات التي تتخذها الشركات استنادا إلى النتائج التي تم الحصول عليها من تحليل البيانات الضخمة، كما يجب على المراجعين الداخليين أن يلعبوا دورا حاسما وهاما في مبادرات البيانات الضخمة الخاصة بمؤسساتهم ، كما يمكن للمراجعين الداخليين الاستفادة من حلول البيانات الضخمة لدعم جهودهم التحليلية للبيانات الخاصة بمشاريع المراجعة الداخلية نظرا لان المنظمة اكتسبت البيانات بالفعل ودمجتها، وبالتالي قد تكتسب المراجعة الداخلية فاعلية كبيرة من خلال استهلاك البيانات من مستودع البيانات أو بحيرة البيانات (DataLake IIA 2017).

ويمكن للبيانات الضخمة تحسين كفاءة تحليلات البيانات الكلية، بما في ذلك التحليلات الوصفية والتشخيصية والتنبؤية والوصفية. يمكن أن توفر هذه التحليلات إحصاءات وصفية عن جميع الافراد، وتقديم أدلة تدقيق على نطاق أوسع وأكثر اكتمالا ، وبناء روابط بين البيانات المالية وعمليات الأعمال الفعلية ، وتحديد الأعلام الحمراء المحتملة (إشارات الإنذار المبكر)، كما يمكن أن تستفيد عمليات التدقيق الداخلي من البيانات الضخمة عن طريق استخدام المزيد من المعلومات وغير المالية للتحكم في المخاطر(,2017Tang&karim) ويرى نيل وايت Neil White **مدير مخاطر ديلويت** "إن القدرة على سحب مصادر البيانات المتباينة معا، وإجراء تحليل بسيط ومعقد، والانضمام إلى مجموعات البيانات، وإلحاقها، والاستفادة من مجموعات أكبر من البيانات الخارجية والمؤسسية هي مهارة حيوية"، ولأن المراجعة الداخلية قد تفتقر إلى الموارد المادية والبشرية من أجل القيام بذلك فلابد من وضع استراتيجية تحليلية برؤية الحالة النهائية التي وضعتها بصورة مشتركة الإدارة العليا والمراجعة الداخلية ،وقد يعقب ذلك تحليل للفجوات لتحديد القدرات المطلوبة من منظور الافراد، والتكنولوجيا والعمليات، ويمكن للفريق بعد ذلك وضع خطة مشروع لمعالجة هذه الثغرات على مدى فترة طويلة، حتى يكون هناك نهج منظم للانتقال من المهارات

الأساسية إلى مهارات تحليلية أكثر تقدما، مع خطة واضحة في مكان وظيفة التدقيق الداخلي يمكن أن تتحرك إلى الأمام في كثير من الأحيان مع سرعة مفاجئة.

ويرى الباحث ان دور المراجع الداخلي لابد أن يتطور ويتكيف مع التطورات التكنولوجية الحديثة والتعامل مع البيانات الضخمة من اجل تزويد الإدارة العليا، ومجلس الإدارة، ولجنة المراجعة والمراجع الخارجي في بعض الأحيان بتقارير واستشارات سريعة ودقيقة من اجل اتخاذ القرارات الصحيحة في الوقت المناسب، وبالتالي يتحتم على المراجعين الداخليين رفع كفاءتهم وفاعليتهم ومستويات تدريبهم من اجل القيام بدورهم في عصر البيانات الضخمة. ولا يمكن إظهار دور المراجع الداخلي في عصر البيانات الضخم الا من خلال توضيح المراحل المختلفة التي تمر بها عمليه تحليل البيانات الضخمة في منظمات الأعمال وهى كما يلى :

المرحلة الأولى: التنقيب عن البيانات

يقوم موفري البيانات الضخمة بجمع البيانات التي تخص منظمات الأعمال من المصادر المختلفة من خلال تطبيقات معينة أو أجهزه استشعار تقوم بالتقاط البيانات بناء على معايير معينة تقوم إدارة الشركة بوضعها من خلال استراتيجية الشركة وأهدافها في التعامل مع البيانات الضخمة يتم الحصول على البيانات وفق نمطين: الفوري أو على دفعات، ففي النمط الفوري، تتدفّق البيانات من المصادر مثل أجهزة الاستشعار أو التطبيقات، بينما يتم جمع البيانات في النمط الثاني من المصادر كل مده معينة أو كمية معينة من البيانات، ويتم نقل البيانات إلى بحيره البيانات Data lake أو موقع مركزي centralized container حتى يتم تشغيلها.

المرحلة الثانية: تحليل البيانات الضخمة

في هذه المرحلة يقوم محللي البيانات الضخمة بتخزين البيانات في بحيرة البيانات Data lake أو موقع مركزي centralized container حتى يتم تشغيلها، ويتم تنظيف البيانات، دمجها، تحويلها وتحليلها وما إلى ذلك وربطها مع بعضها البعض للاستنتاجات هامة ، حتى يتم تقديمها أو عرضها أمام متخذي القرار بالشكل المناسب وفي الوقت المناسب، ويوجد العديد من الأدوات والتقنيات التي تستخدم لتحليل البيانات الكبيرة مثل: Hadoop، MapReduce، GridGain، HPCC، Storm،Cassandr .

المرحلة الثالثة: استخدام البيانات الضخمة

في هذه المرحلة يقوم مختلف المستفيدين ومتخذي القرار داخل منظمات الأعمال مثل مجلس الإدارة، والمدير التنفيذي، ومدير الأقسام المختلفة، والمحاسب الإداري، والمراجع الداخلي وغيرهم بالاستفادة من المعلومات المستخلصة من البيانات الضخمة ويمكن بيان هذه المراحل المختلفة في الشكل التالي.

وينص معيار المراجعة الداخلي 2110 الخاص بالحوكمة على أنه يجب أن يقوم نشاط التدقيق الداخلي بتقييم وتقديم التوصيات المناسبة لتحسين عمليات إدارة المنظمة من أجل:

• اتخاذ القرارات الاستراتيجية والتشغيلية.

• الإشراف على إدارة المخاطر والسيطرة عليها.

• تعزيز الأخلاقيات والقيم المناسبة داخل المنظمة.

• ضمان فعالية إدارة الأداء والمساءلة التنظيمية.

• نقل المخاطر ومعلومات التحكم إلى المناطق المناسبة للمنظمة.

• تنسيق أنشطة، وإيصال المعلومات بين المجلس، المراجعين الخارجيين والداخليين، ومقدمي الضمانات الآخرين والإدارة.

ويمكن دور المراجع الداخلي في كل مرحلة من المراحل السابقة كما يلى:

- **تقييم المخاطر** – هل يمكن أن تساعد التحليلات في تقييم المخاطر المستمر، إما عن طريق الاختبار المخصص أو ضمن عمليات رصد المخاطر المستمرة.
- **التخطيط** – هل يمكن أن تساعد التحليلات على إجراء تدقيق أكثر استهدافا من خلال مساعدة المدققين على التركيز على المناطق أو شرائح السكان الأكثر عرضة للخطر.
- **العمل الميداني** – يمكن استخدام التحليلات لتوفير درجة أعلى من الضمان، أو إجراء الاختبار بشكل أكثر كفاءة.
- **الإبلاغ** – هل يمكن أن نقدم نتائج أكثر ثاقبة نتائج قابلة للتنفيذ من خلال تحليلات، من خلال المساعدة على قياس المخاطر، أو تحديد الأسباب الجذرية.

ويمكن للمراجع الداخلي استخدام المعلومات المستخلصة من تحليل البيانات الضخمة في :

– يمكن أن تساعد البيانات الضخمة المراجعين الداخليين في اكتشاف الفساد والاحتيال داخل منظمات الأعمال، فالاحتيال المالي هو مصدر قلق كبير للمنظمات والاقتصادات في جميع أنحاء العالم، ويقدر اتحاد خبراء الاحتيال المعتمدين Association of Certified Fraud Examiners2016 أن المؤسسة النموذجية تفقد 5 ٪ من العائدات كل عام للاحتيال. وبتطبيق هذا على الناتج العالمي الإجمالي لعام 2014 ، تصل الخسائر العالمية في الاحتيال إلى ما يقرب من 4 تريليون دولار أمريكي. وقد دفعت هذه الأرقام الباحثين إلى النظر في تطبيق تقنيات البيانات الضخمة على كشف الاحتيال والتنبؤ به والوقاية منه. على سبيل المثال ,.Chang et al2008 يقترح استخدام تحليلات البيانات المرئية لفحص ملايين المعاملات البنكية بشكل تفاعلي –فهم يجادلون بأن هذا النهج ممكن وفعال. وعلى النقيض من ذلك، يقوم نموذج ,.Abbasi et al2012 بالاحتيال المالي النموذجي باستخدام "ميتا لينين"meta–leaning ، وهو شكل متخصص في التعلم الآلي يجمع بين مخرجات تقنيات التعلم الآلي المتعددة بطريقة ذاتية التكيف لتحسين الدقة.

أو استخدام الطرق الأخرى الشبكات العصبية الخاضعة للإشراف أو الشبكات العصبية غير الخاضعة للإشراف على أساس خريطة تنظيم ذاتي هرمية متنامية لبناء نموذج كشف الاحتيال المالي (,.Gepp et al 2018).

2– تحسين الاتصال والتواصل اطراف الحوكمة (المراجع الخارجي– لجنة المراجعة– مجلس الإدارة) من خلال إمدادهم بتقارير واستشارات تحتوى على معلومات ذات جودة عالية ناتجة من تحليل بيانات من مصادر متعددة.

3– تفعيل دور المراجع الداخلي في قياس وإدارة الأداء، فالبيانات الكبيرة لديها القدرة على تحسين أنظمة إدارة الأداء من خلال متابعه وقياس ردود أفعال الموظفين ومعنوياتهم والتي قد تظهر من خلال نبرة مشاركات وتعليقات الموظفين في وسائل التواصل الاجتماعي المختلفة ومن خلال اعتماد تقنيات تحليل البيانات الضخمة يمكن تحديد تدابير تحفيزية جديدة والعلاقات بين الأداء الإداري الجيد والمتغيرات التي لم تؤخذ بعين الاعتبار في السابق.

4–المساعدة في تحليل سلوكيات العاملين من خلال تحليل سلوكيات الأفراد على شبكات التواصل الاجتماعي، على سبيل المثال تتضمن النقاشات حول الأمور المتعلقة بالشركات عبر مواقع التواصل معلومات دقيقيه عن سلوكيات العاملين تحمل قيمة كبيرة وبالتالي قد يمتلك العديد من العاملين عده بروفايلات على مواقع التواصل وبجمع هذه البيانات المتعلقة بأحد العاملين من مصادر متعددة ثم

ربطها وتحليلها في الأوقات المناسبة فتعطى إشارات ودلالات هامه عن سلوكيات العاملين ومدى ولائهم للشركة.

وهنا يرى الباحث ضرورة التأكيد على الامور التالية :

- يجب على الجهات المهنية مثل معهد المراجعين الداخليين وغيره، الاهتمام بموضوع البيانات الضخمة وتضمنيها بصوره أساسية ضمن معايير المراجعة الداخلية وفي نفس الوقت تضمين البيانات الضخمة ضمن مقررات الحصول على الشهادات والمؤهلات المهنية.
- يجب على المنظمات التعليمية الاهتمام بشكل اكبر بعلوم البيانات وعلم التنقيب عن البيانات وكيفية الاستفادة من البيانات الضخمة في مجالات الأعمال المختلفة.
- يجب على منظمات الأعمال الاهتمام بالاستفادة القصوى من البيانات الضخمة في اتخاذ مختلف القرارات وفي نفس الوقت يجب ان يتضمن فريق المراجع الداخلية يجب أن يكون مدرب على أحدث وسائل تكنولوجيا المعلومات ا وان يحتوي على مهندس الكترونيات ومتخصص في تحليل البيانات على الأقل.
- يجب أن يمتلك المدققون الداخليون المعرفة والمهارات والكفاءات الأخرى المطلوبة أداء مسؤولياتها الفردية. يجب أن يمتلك نشاط التدقيق الداخلي بشكل جماعي أو الحصول على المعرفة والمهارات والكفاءات الأخرى اللازمة لأداء مسؤولياتها. وبالتالي يجب على فريق المراجعة الداخلية ان يكون عنده من المهارات التي تمكنه من التعامل مع البيانات الضخمة من مرحله توفير البيانات الضخمة وتحليليها والاستفادة منها في اتخاذ قرارات المراجعة الداخلية (حسب نص المعيار المراجعة الداخلية 1210) .

البحث الثاني : الرقابة الحكومية للبيانات الضخمة (Big Data)

الحكومات أيا كانت من أي بلد يوجد لديها كمية هائلة جدا من البيانات على أساس يومي تقريبا و السبب هو لديهم أجهزة لتتبع السجلات وقواعد البيانات المتعلقة المواطنين. يمكن أن يكون للبيانات الضخمة تأثير هائل في الحكومة، فإن الحكومات تحاول لجعل جميع المعلومات التي تتلقاها ذات مغذى ومساعدتها في اتخاذ

القرارات الحيوية التي تؤثر على الملايين من الناس. ومن الأمثلة على ذلك: تحليل البيانات الضخمة لعب دوراً كبيراً في حملة إعادة انتخاب الرئيس الأمريكي باراك أوباما عام 2012

وعندما يتعلق الأمر بإدارة البيانات، أغلب المنظمات الحكومية تواجه مشكلة وجود كميات هائلة من البيانات في أنظمة الكمبيوتر، ومعظم هذه البيانات غير منظمة أو مُهيكلة (unstructured data) وهذا يعني أنها لا تناسب أي نموذج بيانات معرّف مسبقاً. لفهم الأنماط الموجودة في هذه البيانات يجب أن تطبق المنظمات الحكومية نماذج إحصائية تسعى لالتقاط ومعالجة كميات هائلة من البيانات غير المهيكلة و تسمى هذه العملية بالبيانات الضخمة .وأغلب المنظمات الحكومية لا تملك عدد كافي من الموظفين أو لا تملك القدرة الحسابية اللازمة لإدارة وتحليل جميع البيانات الخاصة بهم، ومع طبيعة البيانات المتغيرة وارتفاع حجمها أصبحت الاستعانة بأدوات البيانات الضخمة من خلال الحوسبة السحابية (Cloud Computing) أمراً ضرورياً. فأصبح بإمكان المختصين بتطوير الخدمات الحكومية رصد مدى رضا المواطنين عن الخدمات المقدَّمة لهم. وعلى ضوء النتائج المحللة يمكن استنتاج ما يلزم عمله للتطوير والتحسين، حيث أصبح مسح آراء الجمهور عن طريق الاستبيانات التقليدية مكلفاً وغير مجدٍ في كثير من الأحيان، وذلك نظراً لتنوع البيانات الديموغرافية وثقافات المتعاملين. إن من أكبر المصادر لتلك البيانات الضخمة هي البيانات المسجلة من خلال عمليات التعداد السكاني والتسجيل في قواعد البيانات الحكومية، حيث يمكن أن تستنتج الحكومات معلومات ثمينة جداً من خلال تحليل تلك البيانات المخزنة.

<u>يجب على الأجهزة العليا للرقابة [13]وضع البيانات على جدول الاعمال:</u>

- تقييم القدرات الداخلية (الافراد والعمليات والتكنولوجيا والاستراتيجيات) لتصبح أكثر اعتمادًا على البيانات .
- نظرة متفائلة حول جاهزية الأجهزة العليا للرقابة لتحليل البيانات الضخمة .
- تقييم كيفية استخلاص القيمة من تحليلات البيانات .
- بناء الشراكات لضمان الوصول إلى البيانات ذات الصلة .

[13]العرض المقدم في اللقاء العلمي حول البيانات الضخمة وتأثيرها على التقارير.2020 يناير 18 من مدير تدقيق نظم المعلومات بالمنظمة الإفريقية للأجهزة العليا للرقابة الناطقة باللغة الانكليزية "فروساي –أي " منتدب من مكتب المراجع العام في زامبيا.

– القيام بتمكين الموظفين ووضع خطة لكيفية التوظيف مستقبلا .

– بناء القدرات والبنية التحتية.

وذلك من خلال الافراد وتحديد واضح للأدوار والمسؤوليات و مهارات البيانات داخليا وخارجيا و البرامج التدريبية الصحيحة او الملائمة و المدققون و خبراء تحليل البيانات.

وهنا لابد من التركيز على مواضيع مثل :

- مدى تطبيق الأجهزة الحكومية تقنيات وبرامج تحليل البيانات الضخمة
- مدى تطبيق الأجهزة الرقابية تقنيات وبرامج تحليل البيانات الضخمة على عملية الرقابة وجودتها
- دور البيانات الضخمة في اكتشاف الغش في الشركات الحكومية
- أثر البيانات الضخمة على تطوير مقاييس الأداء في الشركات الحكومية
- أثر البيانات الضخمة على ادله الأثبات في الرقابة الحكومية
- اثر البيانات الضخمة على اختبارات الرقابة واختيار العينات والاهمية النسبية وإعداد التقرير.
- تم تأسيس مجموعة عمل الانتوساي للبيانات الضخمة في مؤتمر INCOSAI تُسهل مجموعه عمل البيانات الضخمة تبادل المعلومات والخبرات بين الأجهزة العليا للرقابة، والتي بدورها تُسلح الأجهزة العليا للرقابة بشكل أفضل للمساهمة في الحوكمة الرشيدة والتنمية محلياً وعالمياً المستدامة
- في عام2018 ،أصدرت مجموعة عمل البيانات الضخمة مجموعة من حالات مراجعة البيانات الضخمة العالمية، "الممارسات العالمية: مراجعه البيانات الضخمة من أجل التنمية المستدامة"، والتي تعكس إلى أي مدى يمكن لتبادل الخبرات تعزيز التعاون، وزيادة الوعي ودفع التقدم في مجال مراجعة البيانات الضخمة.

وبمكن تلخيص مهمة مجموعة عمل البيانات الضخمة بما يلي :

- تلخيص وتعزيز ممارسات وخبرات مراجعة البيانات الضخمة.
- إصدار إرشادات لمراجعة البيانات الضخمة وتقارير الأبحاث.
- دعم التطوير المهني لمراجعي البيانات الضخمة.
- تعزيز التعاون الثنائي ومتعدد الأطراف بين الأجهزة العليا للرقابة.
- تسهيل أنشطة تبادل المعرفة بالبيانات الضخمة.

- تقديم مقترحات عملية تتناول مقدمة أو تحسين القوانين
- تقديم خطط قصيرة، متوسطة وطويلة الأجل لمراجعة البيانات الضخمة بما يتوافق مع الظروف الوطنية؛ ونماذج التنظيم والإدارة المناسبة لمراجعة البيانات الضخمة .

- وفي المؤتمر الدولي السنوي للنزاهة 2019 الذي استضافه الجهاز الأعلى للرقابة ببيرو تحت عنون "استخدام التكنولوجيا لمكافحة الفساد " وناقش المؤتمر قضايا الذكاء الاصطناعي والبيانات الضخمة وذكاء الأعمال والتعلم الآلي واثرها على تسهيل اكتشاف ومنع أفعال سوء السلوك الوظيفي والفساد في الإدارة العامة على النطاق العالمي.

- في عام 2015 قامت غرفة الحسابات الروسية بإجراء عمليات المراجعة عن بُعد باستخدام نظم تحليل المعلومات المتكاملة و تطوير وادخال نظام لتحليل المعلومات الجراء عمليات المراجعة عن بُعد استنادًا إلى القانون الفيدرالي لعام 2013 الذي يمنح الجهاز الأعلى للرقابة صالحية الوصول مباشرة إلى أنظمة معلومات الجهات الخاضعة للمراجعة[14]. نظراً لاستخدام المؤسسات الحكومية بشكل متزايد لتكنولوجيا المعلومات لتحقيق الأهداف، فقد تقوم الأجهزة العليا للرقابة بعمليات المراجعة عن بعد كبديل لعمليات المراجعة التقليدية في الموقع (المراجعات الميدانية).

إن التنفيذ الناجح لغرفة الحسابات لعمليات المراجعة عن بعد باستخدام نظام تحليل المعلومات يوحي بأن اجهزة المراجعة في البلدان الأخرى قد تستفيد من نهج مماثل. وعلى الرغم من أهمية الاستمرار في تحسين لقد أثر توافر البيانات وتحليلات البيانات في المراجعة – كمجال ومهنة – خاصة وأن رقمنة العملية المالية جلبت تقنيات جديدة تسمح للأجهزة العليا للرقابة بمراجعة 100 ٪من المعاملات. تزيد هذه القدرة من كفاءة المراجعة المالية وتتيح للمراجعين التمييز بين المعاملات العادية وغير العادية. إن تمييز المعاملات غير العادية يخلق كفاءة إضافية، حيث لم يعد المراجعون بحاجة إلى مراجعة عينة كبيرة .

وتركز مراجعة الأداء تقليديًا على تحديد مدى اقتصادية وكفاءه البرنامج. ومع ذلك، فإن البيانات المتاحة ووسائل تحليل كميات كبيره منها بشكل سهل ويسير تسمح للمراجعين بقياس فاعلية البرنامج مباشرة.

على سبيل المثال، عند **مراجعة برنامج الحد من الفقر**، يمكن للأجهزة العليا للرقابة الآن تقييم التغيرات الفعلية في الفقر نتيجة تدخل الحكومة. و قد يتم الحصول على نسب المشاركة في سوق العمل من قواعد البيانات الضريبية التي يمكن أن تفتقر الى معلومات هامة اللازمة للمراجعة.

[14] مقال بقلم ميخائيل بيتروف، مدير إدارة التحول الرقمي غرفة الحسابات بالاتحاد الروسي .

قبل إجراء التقييمات باستخدام المعلومات من قواعد الاعتبار مبكراً ان وضع الحاجة الى برامج ضرورية في له فوائد عديدة: يحتمل أن يخفض التكاليف عند تلبية احتياجات المراجعة الحالية والمستقبلية؛ توفير معيار للمهارات؛ واختيار حلول مناسبة لحماية البيانات. وقد تتطلب حلول البرامج الجديدة أي أيضًا أجهزة حاسب الي عالية الأداء – لكل محلل بيانات و / أو للخوادم المركزية المشتركة أو حلول معتمدة على نظم السحابة. قد تعتمد هذه الخيارات أي أيضا على كيفية تخصيص متخصصي البيانات. إذا كان اختصاصيو البيانات، على سبيل المثال، يعملون على عمليات المراجعة من البداية إلى النهاية، فقد يكون من المنطقي تجميع موارد الحواسب مع الخوادم المركزية أو المشتركة أو الحلول المعتمدة على تقنية السحابة، حيث سيكون عدد المتخصصين الذين يقومون بتحليل البيانات محدودًا. حيث اختار الجهاز الأعلى للرقابة ببلجيكا تخصيص متخصصي البيانات لفرق المراجعة، حيث يقوموا بإجراء عمليات المراجعة من البداية إلى النهاية. ومع ذلك، ومن خلال مجموعة المناقشة DataLab ، يقوم متخصصو البيانات أيضاً بدعم الفرق التي ليس لديها متخصصون في تحليل البيانات.

الفصل الثالث دراسة ميدانية – الجهاز المركزي للرقابة المالية – سورية

الجهاز المركزي للرقابة المالية هو هيئة رقابية مستقلة ترتبط برئيس مجلس الوزراء وتهدف أساساً إلى تحقيق رقابة فعالة على أموال الدولة ومتابعة أداء الأجهزة التنفيذية الإدارية والاقتصادية لمسؤولياتها من الناحية المالية ويختص بتدقيق وتفتيش حساباتها وذلك على الوجه المبين في المرسوم التشريعي 64 لعام 2013[15] ويقوم في مجال الرقابة المالية والمحاسبية بما يلي :

- مراجعة الحسابات والميزانيات والقوائم المالية الختامية لمؤسسات وشركات ومنشآت القطاع العام الاقتصادي للتعرف على حقيقة المركز المالي وفقاً للمبادئ المحاسبية السليمة ومعايير المحاسبة والمراجعة المعتمدة وإبداء الملاحظات بشأن الأخطاء والمخالفات والقصور في تطبيق القوانين والأنظمة.
- رقابة الكفاية والأداء والتحقق من أن استخدام الموارد المالية قد تم بأعلى درجة من الكفاية دون إسراف أو ضياع فضلاً عن مراقبة مدى كفاية استعمال وإدارة الأموال الموضوعة تحت تصرف هذه الجهات العامة ومراجعة حسابات تكاليف الأعمال على ما كان مقدراً لها وتقييم النتائج ومقارنتها بالاستثمارات وتكلفتها والموارد المستخدمة فيها والتحقق من أن تلك الأعمال قد تم تنفيذها بالمصروفات التي قدرت لها دون إسراف أو ضياع أو تبذير.

[15] - المرسوم التشريعي رقم / 64 / لعام 2013 المتضمن قانون الجهاز المركزي للرقابة المالية

الامر الذي جعل الجهاز المركزي في سورية طرف ذو مصلحة في التحول الرقمي وعضو فاعل في بناء قاعدة البيانات للحكومة الإلكترونية بناء عليه قام بإعداد و تنظم مصفوفة بيانات ضخمة لمواضيع رقابية مثل" مصفوفة العقود " لكافة الوزرات والجهات العامة. وكذلك اعتماد استراتيجية تدقيق أهداف التنمية المستدامة المعتمدة عالمياً إضافة إلى مهامه الرقابية الاعتيادية مما دفعه للمساهمة الجدية في استخدام التقنيات الحديثة لتحليل البيانات الضخمة .

-تم اجراء الدراسة الميدانية على عينة من موظفي الجهاز المركزي للرقابة المالية في مدينة دمشق، حيث بلغ حجم المجتمع المدروس (183) مفتش ومفتشة من مختلف المستويات الإدارية، بهدف التعرف على تقنيات تحليل ورقابة البيانات الضخمة من خلال أنواعها وخصائصها، وتسليط الضوء على الطرق والأدوات المستخدمة لتحليل البيانات الضخمة، وأهم القضايا المتعلقة بها.

ولتحديد حجم العينة المدروسة قام الباحث باستخدام جداول كريجسي ومورغان حيث أن حجم العينة المناسب لحجم مجتمع (183) هو (123)، ومن أجل ذلك تم توزيع (130) استبيان تم استرداد وإلغاء عدد منها، كما يبينه الجدول التالي:

الجدول رقم (1) يبين عدد الاستبيانات الموزعة والمستردة والصالحة للتحليل على أفراد عينة الدراسة

البيان	العدد	النسبة
الاستبيانات الموزعة	130	100%
الاستبيانات المستردة	122	94%
الاستبيانات غير المستردة	8	6%
الاستبيانات غير الصالحة للتحليل	5	4%
الاستبيانات الصالحة للتحليل	117	90%

نلاحظ من الجدول رقم (1) أن عدد الاستبانات المستردة والقابلة للتحليل بلغت (117) استبانة بنسبة 90% من الاستبانات الموزعة وهي نسبة كافية ومقبولة لتحليل البيانات، وبذلك يستقر العدد الكلي للأفراد العينة على 117 فرداً، حيث كانت العينة تتوزع وفق المتغيرات الديموغرافية كما يبينه الجدول رقم (2):

الجدول رقم (2) يبين التوزيع التكراري والنسبي لعينة البحث وفقاً للمتغيرات الديمغرافية

المتغيرات الديموغرافية		العدد	النسبة المئوية
الجنس	أنثى	32	27.35%
	ذكر	85	72.65%
	المجموع	117	100.00%
الدرجة العلمية:	إجازة جامعية	67	57.26%
	دبلوم عالي	4	3.42%
	درجة ماجستير	42	35.90%
	درجة دكتوراه	4	3.42%
	المجموع	117	100.00%
سنوات الخبرة	أقل من 10 سنوات	38	32.48%
	أكثر من 10 سنوات وأقل من 15 سنوات	22	18.80%
	من 15 سنة وأكثر	57	48.72%
	المجموع	117	100.00%
المنصب الوظيفي	مدير	27	23.08%
	مفتش	20	17.09%
	مفتش أول	58	49.57%
	مفتش معاون	12	10.26%
	المجموع	117	100.00%

وبتحليل المتغيرات الديموغرافية لعينة الدراسة نلاحظ أن أغلب أفراد العينة كانوا حاصلين على إجازة جامعية حيث بلغت نسبتهم (57.26%) من العينة، في حين بلغت نسبة حملة شهادة الماجستير (35.9%)، أما فيما يتعلق بخبرة أفراد العينة فنجد أن (32.48%) خبرتهم أقل من 10 سنوات، و(18.8%) تتراوح خبرتهم بين 10 و15 سنة، أما بالنسبة للمنصب الوظيفي لأفراد العينة يتبين أن (49.57%) من أفراد العينة يشغل وظيفة مفتش أول، و(23.08%) مدير، و(17.09%) بمنصب مفتش أول، و(10.26%) مفتش معاون.

1. تصميم الإستبيان واختبار صدق المقاييس وثباته:

1.3. تصميم الاستبانة:

قام الباحث بتصميم استمارة استبيان تغطي الجوانب المتعلقة البيانات الضخمة، وقد قُسمت إلى خمسة أقسام أساسية، ضم القسم الأول بعض البيانات الديموغرافية المتعلقة بأفراد عينة البحث كالمؤهل العلمي، وعدد سنوات الخبرة، والموقع الوظيفي.

أما القسم الثاني فقد تكون من (5) عبارات تقيس إدراك العاملين للمعرفة بالبيانات الضخمة.

القسم الثالث فقد تكون من (5) عبارات تقيس تحليل البيانات الضخمة.

والقسم الرابع فقد تكون من (10) عبارات تقيس معرفة العاملين بتقنيات تحليل البيانات الضخمة.

القسم الخامس فقد تكون من (10) عبارات تقيس رقابة البيانات الضخمة من قبل الجهاز المركزي للرقابة المالية في الجمهورية العربية السورية

ولذلك استخدام الباحث مقياس ليكرت لقياس استجابات أفراد العينة على فقرات الاستبيان المكون من خمس رتب تراوحت بين موافق بشدة وغير موافق بشدة، وقد صحح المقياس وفقاً للدرجات الآتية: (5 موافق بشدة، 4 موافق، 3 محايد، 2 غير موافق، 1 غير موافق بشدة).

3.2. اختبار صدق المقاييس وثباته:

يقصد بصدق المقياس مدى قدرته على قياس الشيء المراد قياسه بدقة ومن أجل ذلك عرض الباحث الإستبانة على مجموعة من المحكمين المتخصصين في المجالات الرقابية وتحليل البيانات، حيث أشاروا إلى بعض الملاحظات كحذف أو تعديل بعض العبارات وقد استجاب الباحث لآراء المحكمين وقام بإجراء ما يلزم ضوء المقترحات المقدمة، وبذلك خرج الاستبيان في صورته النهائية. أما ثبات المقياس فُيقصد به الاتساق الداخلي بين عبارات المقياس، وتوجد عدة طرائق لحساب ثبات المقاييس إلا أن الباحث اعتمد على طريقة ألفا كرونباخ Cronbach's Alpha لحساب ثبات المقياس باستخدام البرنامج الإحصائي SPSS وقيمة معامل الاتساق ألفا كرونباخ يجب أن تتراوح بين 0 و 1 وحتى يتمتع المقياس بالثبات يجب ألاّ يقل الحد الأدنى لقيمة معامل الاتساق ألفا كرونباخ في هذا الاختبار عن 0.6.

أشارت النتائج إلى أن قيمة ألفا كرونباخ بلغت (0.731) بالنسبة إلى متغير المعرفة بالبيانات الضخمة، و(0.785) بالنسبة إلى متغير تحليل البيانات الضخمة، و(0.720) بالنسبة إلى متغير تقنيات التحليل البيانات الضخمة، و(0.740) للرقابة البيانات الضخمة من قبل الجهاز المركزي للرقابة المالية في الجمهورية العربية السورية، هو موضح بالجدول رقم (3)، وتعد هذه القيم مقبولة في البحوث الرقابية والاقتصادية والمالية، وبناءً على ذلك يمكن القول: إن المقاييس المستخدمة جميعها في البحث تتسم بالثبات الداخلي لعباراتها.

الجدول رقم (3) معاملات ألفا كرونباخ لمتغيرات البحث

المتغير	معامل ألفا كرونباخ	عدد البنود
البيانات الضخمة	0.731	5
تحليل البيانات الضخمة	0.785	5
كفاية وملائمة تقنيات تحليل البيانات الضخمة	0.720	10
رقابة البيانات الضخمة من قبل الجهاز المركزي للرقابة المالية في الجمهورية العربية السورية	0.740	10

2. دراسة الفرضيات:

1.2. الفرضية الأولى: يتوافر لدى المدققين والاجهزة الحكومية معرفة جيدة بماهية البيانات الضخمة عند مستوى دلالة إحصائية (0.05).

لدراسة الفرضية الأولى قام الباحث بحساب المتوسط الحسابي والانحراف المعياري لمتغير المعرفة بالبيانات الضخمة، ثم تمت مقارنة متوسط مع القيمة المتوسطة (3) باستخدام اختبار ت ستيودنت (T-Test) كما يبينه الجدول التالي:

الجدول رقم (4) يبين المتوسط الحسابي والانحراف المعياري ونتائج اختبار ت ستيودنت (T-Test)

	المعرفة بالبيانات الضخمة
العدد	117
المتوسط الحسابي	3.7248
الانحراف المعياري	0.49600
t	15.806
df	116
Sig. (2-tailed)	0.001

من الجدول السابق نجد أن قيمة المعنوية الإحصائية (Sig = 0.001) أصغر من مستوى الدلالة الإحصائية (0.05) وبالتالي نقبل الفرضية الأولى أي يتوافر لدى المدققين والاجهزة الحكومية معرفة جيدة بماهية البيانات الضخمة عند مستوى دلالة إحصائية (0.05) لأن قيمة المتوسط الحسابي (3.7248) أكبر من القيمة المتوسطة.

2.2. الفرضية الثانية: يستخدم الجهاز الاعلى للرقابة تحليل البيانات اثناء عملية التدقيق عند مستوى دلالة إحصائية (0.05).

لدراسة الفرضية الثانية قام الباحث بحساب المتوسط الحسابي والانحراف المعياري لمتغير تحليل البيانات الضخمة، ثم تمت مقارنة متوسط مع القيمة المتوسطة (3) باستخدام اختبار ت ستيودنت (T-Test) كما يبينه الجدول التالي:

الجدول رقم (5) يبين المتوسط الحسابي والانحراف المعياري ونتائج اختبار ت ستيودنت (T-Test)

تحليل البيانات	العدد	المتوسط الحسابي	الانحراف المعياري	t	df	Sig. (2-tailed)
الضخمة	117	2.2134	0.51203	-16.575	116	0.001

من الجدول السابق نجد أن قيمة المعنوية الإحصائية (Sig = 0.001) أصغر من مستوى الدلالة الإحصائية (0.05) وبالتالي نرفض الفرضية الثانية أي لا يستخدم الجهاز الاعلى للرقابة تحليل البيانات اثناء عملية التدقيق عند مستوى دلالة إحصائية (0.05) لأن قيمة المتوسط الحسابي (2.2134) أقل من القيمة المتوسطة.

2.2. الفرضية الثالثة: يتوفر لدى الجهاز المركزي الكفاية والملائمة لتحليل البيانات الضخمة عند مستوى دلالة إحصائية (0.05).

لدراسة الفرضية الثالثة قام الباحث بحساب المتوسط الحسابي والانحراف المعياري لمتغير كفاية وملائمة للتقنيات التحليلية، ثم تمت مقارنة متوسط مع القيمة المتوسطة (3) باستخدام اختبار ت ستيودنت (T-Test) كما يبينه الجدول التالي:

الجدول رقم (6) يبين المتوسط الحسابي والانحراف المعياري ونتائج اختبار ت ستيودنت (T-Test)

كفاية وملائمة تحليل البيانات الضخمة	العدد	المتوسط الحسابي	الانحراف المعياري	t	df	Sig. (2-tailed)
	117	3.217	0.580	4.052	116	0.001

من الجدول السابق نجد أن قيمة المعنوية الإحصائية (Sig = 0.001) أصغر من مستوى الدلالة الإحصائية (0.05) وبالتالي نقبل الفرضية الثالثة أي يتوفر لدى الجهاز الاعلى للرقابة المالية الكفاية

والملائمة لمعالجة البيانات الضخمة عند مستوى دلالة إحصائية (0.05) بشكل متوسط لأن قيمة المتوسط الحسابي (3.217).

2.2. الفرضية الرابعة: ما مدى تطبيق تقنيات تحليل البيانات الضخمة لدى أجهزة الرقابة (الداخلية والحكومية) والاستفادة منها في خلق قيمة مضافة وجودة وموثوقية للعملية الرقابية عبر مراحلها المتعددة عند مستوى دلالة إحصائية (0.05).

لدراسة الفرضية الرابعة قام الباحث بحساب المتوسط الحسابي والانحراف المعياري لمتغير تطبيق تقنيات تحليل البيانات، ثم تمت مقارنة متوسط مع القيمة المتوسطة (3) باستخدام اختبار ت ستيودنت (T-Test) كما يبينه الجدول التالي:

الجدول رقم (7) يبين المتوسط الحسابي والانحراف المعياري ونتائج اختبار ت ستيودنت (T-Test)

تقنيات تحليل البيانات الضخمة	العدد	المتوسط الحسابي	الانحراف المعياري	t	df	Sig. (2-tailed)
	117	2.9000	0.54314	-1.992	116	0.049

من الجدول السابق نجد أن قيمة المعنوية الإحصائية (Sig = 0.049) أصغر من مستوى الدلالة الإحصائية (0.05) وبالتالي نقبل الفرضية الرابعة أي يتوفر لدى الجهاز الاعلى للرقابة المالية تقنيات تحليل البيانات الضخمة عند مستوى دلالة إحصائية (0.05) بشكل متوسط لأن قيمة المتوسط الحسابي (2.9).

ومما سبق نجد ما يلي:

- أهمية البيانات ضخمة في تخزينها أو تحليلها من قبل الأجهزة والبرمجيات التقليدية .وحجم البيانات المستخرجة من مصدر ما، هو ما يحدد قيمة وإمكانيات البيانات لكي تُصنف من ضمن البيانات الضخمة ومن مزايا استخدام أداة تحليل البيانات تحسين القدرة على تحديد المخاطر وتخفيفها كما أن الحصول على البيانات الضخمة يتطلب القدرة على الحصول على البيانات خاصة من المصادر النوعية .

- هنالك دور هام لتحليل البيانات وتقنيات الرقابة والذكاء الاصطناعي والتعلم الآلي على تحليل كمياتٍ كبيرةٍ من البيانات بالاعتماد على الخوارزميات المتقدمة.

– معالجة البيانات الضخمة للمنظمات وتحليلها يمكن التوصل إلى مستويات متقدمة جدا من الذكاء التنبؤي والقدرة على فهم نمط وسلوكيات ورغبات الأشخاص .

– تعتمد الحكومات على تقنيات الذكاء الاصطناعي التي لديها القدرة على التعلم والاستنتاج ورد الفعل لتحليل البيانات ومعالجتها .

– في تقنيات تحليل البيانات الضخمة لا يوجد تطبيق كافي وملائم لتحليل البيانات الضخمة في القطاع العام يساعد الجهاز الاعلى للرقابة على تحليل البيانات اثناء عملية التدقيق وكما أنه تتوفر تقنيات تحليل البيانات الضخمة لدى أجهزة الرقابة خبرات فنية قادرة على القيام بعمليات تحليل البيانات الضخمة .

– لا يستفاد من تحليل الضخمة خلق قيمة مضافة وجودة وموثوقية للعملية الرقابية عبر مرحلها المتعددة

– تؤثر البيانات الضخمة على اختبارات الرقابة واختيار العينات والاهمية النسبية وإعداد التقرير

– هنالك دور وأثر لتحليل البيانات الضخمة في اكتشاف الغش ولاحتيال في الجهات الحكومية وكذلك على أدله الأثبات في الرقابة الحكومية .

–لا يتوافر لدى الجهاز العناصر التالية (الأفراد، العمليات ،التكنولوجيا) للقيام بعملية التحليل البيانات الضخمة

– تحليل البيانات من قبل الأجهزة العليا للرقابة تضيف قيمة للتقارير والمواطنين.

–ليست كامل العمليات الحكومية الرئيسية مؤتمتة ولا تعتبر موفر خدمة موثوق للقيام بتحليل البيانات الضخمة

– الرقابة على البيانات الضخمة تمكن من قياس وتقييم الاداء بظل عمل الحكومة الإلكترونية و التحول الرقمى .

– يوجد استخدام استراتيجي للبيانات الضخمة في اتخاذ القرارات الرئيسية وتقديم الخدمات

النتائج والتوصيات

• البيانات الضخمة تزداد ضخامة بسرعة هائلة وتحتاج إلى خطة معالجة استراتيجية .

• الفائدة من البيانات الضخمة لا زالت محدودة مقارنة بما فيها من فرص ونتائج يمكن الحصول عليها بعد التحليل والقيمة المضافة لهذه البيانات عندما يتم تحليلها ومعالجتها.

–لا توجد كفاية وملائمة لتقنيات التحليل الاحصائية والبرمجية المستخدمة لمعالجة البيانات الضخمة سواء لدى الحكومة أو الاجهزة الرقابية .

– تطبيق تقنيات تحليل البيانات الضخمة لدى أجهزة الرقابة (الداخلية و الحكومية) والاستفادة منها في خلق قيمة مضافة وجودة و موثوقية للعملية الرقابية عبر مرحلها المتعددة لازالت حديثة الولادة .

– أن تقنيات الحوسبة السحابية من أهم الحلول والطرق المفيدة في ايجاد مساحات تخزينية للبيانات الضخمة وتساعد في التعامل معها بالسرعة المناسبة وأوصت الدراسة بضرورة سن الانظمة والتشريعات المنظمة على المستوى الوطني والإقليمي والدولي بشكل واضح لموضوع الملكية الفكرية وخصوصية البيانات.

– استثمار البرمجيات مفتوحة المصدر المتخصصة في تحليل البيانات والمواءمة بين المناسب منها ومحطات العمل اللازمة".

– وضع استراتيجيات تساعد على إدارة البيانات الضخمة بكامل التفاصيل التي تدعم اتخاذ القرار.

– يفيد تحليل البيانات في تبويبها بشكل يسهل فهمها و أيضا تفيد في تسليط الضوء على نقاط القوة التي يمكن استعمالها للتغلب على المنافسين و تسليط الضوء على نقاط الضعف التي يمكن ابتكار عدة حلول لمعالجتها

–إمكانية التكوين في مجال البيانات حتى يتم تكوين أشخاص من ذوي الخبرة، مدربين على التعامل بكفاءة مع البيانات ويتحكمون في كيفية استغلالها في شتى قطاعات التنمية.

– وضع استراتيجية تخص تعزيز البيانات الضخمة، وتخدم المصلحة العامة، وتحديد الجهات، والمؤسسات المسؤولة عن جمع وتحليل، ومعالجة البيانات تعمل على تطبيق معايير دولية موحدة للتعامل مع البيانات.

– تشجيع وتعزيز التعاون الدولي، والاستفادة من خبرات المختصين في معالجة وتسيير البيانات.

– ضع آليات شراكة عامة وخاصة من شأنها تسهيل نقل المعرفة، وتبادل البيانات الجديدة المنبثقة من البحث والتطوير، علاوة على الإبداع في إنتاج الاحصاءات الرسمية.

التوصيات :

- طالما هناك دور هام لتحليل البيانات وتقنيات الرقابة والذكاء الاصطناعي والتعلم الآلي على تحليل كمياتٍ كبيرةٍ من البيانات بالاعتماد على الخوارزميات المتقدمة وجب التعرف عليها .
- دعم الحكومات على تطبيق تقنيات الذكاء الاصطناعي والتعلم الالي واستخدامها عند التحول الرقمي.

- تطبيق كافي وملائم لتحليل البيانات الضخمة في القطاع العام يساعد الجهاز الاعلى للرقابة على تحليل البيانات اثناء عملية التدقيق وكما أنه تتوفر تقنيات تحليل البيانات الضخمة لدى أجهزة الرقابة خبرات فنية قادرة على القيام بعمليات تحليل البيانات الضخمة .

- تحليل البيانات الضخمة لدى الاجهزة الرقابية يساعد في اكتشاف الغش .

- ضرورة توفير مستلزمات رقابة البيانات الضخمة من (الأفراد، العمليات ،التكنولوجيا) .

- تحليل البيانات من قبل الأجهزة العليا للرقابة تضيف قيمة للتقارير والمواطنين.

- لابد من استخدام استراتيجي للبيانات الضخمة في اتخاذ القرارات الرئيسية وتقديم الخدمات
- نوصى بأجراء المزيد من البحوث المتعلقة بكيفية الاستفادة من البيانات الضخمة واستحداث طرق جديده لتحليل البيانات الضخمة في مجال المحاسبة والمراجعة والقيام بمجموعه من الدراسات المستقبلية مثل :
  - دور البيانات الضخمة في اكتشاف الغش في الشركات الحكومية
  - أثر البيانات الضخمة على تطوير مقاييس الأداء في الشركات الحكومية
  - أثر البيانات الضخمة على ادله الأثبات في المراجعة.

المراجع

1- الامم المتحدة، المجلس الاقتصادي والاجتماعي، اللجنة الاحصائية، البيانات الضخمة وتحديث النظم الاحصائية، مارس 2014 ،

2- موقع سين نيوز 2021 – ما هي تحليلات البيانات الكبيرة؟

3- أثر تحليل البيانات الضخمة علي تحسين جودة المعلومات المحاسبية –د. نجاة محمد مرعى يونس أستاذ المحاسبة المساعد بكلية الأعمال جامعة جدة.

4- البيانات– الضخمة و مجالات تطبيقها – إعداد د /عدنان مصطفى البار– أستاذ نظم المعلومات المشارك، كلية الحاسبات وتقنية المعلومات، جامعة الملك عبدالعزيز

5- الادوار الجديدة لأخصائي المعلومات للتعامل مع البيانات الضخمة – د/ زينب بن الطيب 1 و د/ سليمان بن إبراهيم الرباعي – تاريخ النشر الكترونيا: 10 يناير 2019 –

6- البيانات الضخمة واتخاذ القرار في جامعة الملك سعود: دراسة تقييمية– لنظام اتقان – د. علي بن ذيب الأكلبي – جامعة الملك سعود – نائب المشرف لشؤون النشر العلمي

7- مقال : أ.د. أحمد صادق عبد المجيد– أستاذ تكنولوجيا التعليم بكلية التربية جامعة سوهاج

8- دور المراجع الداخلي في عصر البيانات الضخمة Big Data(الجزء الثانى) –حماده السعيد المعصراوي – باحث دكتوراه

9- المقال منشور فى العدد 79 من مجلة الاقتصاد الإسلامي العالمية GIEM تاريخ النشر الاثنين، 31 ديسمبر 2018

10- العرض المقدم في اللقاء العلمي حول البيانات الضخمة وتأثيرها على التقارير.2020 يناير 18 مدير تدقيق نظم المعلومات بالمنظمة الإفريقية للأجهزة العليا للرقابة الناطقة باللغة الانكليزية "فروساي –أي " منتدب من مكتب المراجع العام في زامبيا.

11- اسلام النجار 11 أكتوبر، 2020 – – الذكاء الاصطناعي وعلم البيانات..

12- افضل 10 أدوات لتحليل البيانات – WINDOWS– A By 7 Home0la Home دليل التقنيةWindows

13- نموذج مبادرة تنمية الانتوساي للتدقيق على أهداف التنمية المستدامة isam – النسخة الاسترشادية – 2020

14- معيار المراجعة الداخلي 2110 الخاص بالحوكمة .

15- أ.د. أحمد صادق عبد المجيد ثورة البيانات الضخمة - عالم الثقافة نوفمبر 14, 2020 - أستاذ تكنولوجيا التعليم بكلية التربية جامعة سوهاج.

16- مقال بقلم ميخائيل بيتروف، مدير إدارة التحول الرقمي غرفة الحسابات بالاتحاد الروسي .

17- المرسوم التشريعي رقم / 64 / لعام 2013 المتضمن قانون الجهاز المركزي للرقابة المالية.

المراجع الاجنبية

- Stryk B. 2015;, do organizations prepare and clean Big Data to achieve better data governance? A Delphi Study (doctoral dissertation, Capella University
- Janssen M, van der Voort H, Wahyudi A. 2017;, Factors influencing big data decision-making quality. Journal of Business Research, 70 (2017) 338-345. DOI: https://doi.org/10.1016/j.jbusres.2016.08.007 .
- Data Management Association International (DAMA) • Data Science Association • Digital Analytics Association (DAA) • International Institute for Analytics (IIA) • International Machine Learning Society (IMLS) • Institute for Operations Research and the Management Sciences (INFORMS) • SIGKDD
- Maaroof, A. (2015). Big Data and the 2030agenda for sustainable development. the Development Goals in the Asia and the Pacific region, (pp. 52-1). Bangkok 2019 .Author(s), licensee HBKU Press.

إعداد الباحث : زياد ابراهيم الخطيب