جامعة الملك سعود
كلية الآداب
مجلة الآداب
ردمد (ورقي): ٣٦٢٠-١٠١٨
ردمد (النشر الإلكتروني): ٨٣٣٩-١٦٥٨

King Saud University
College of Arts
Journal of Arts
ISSN (Paper):1018-3620
ISSN (Electronic):1658-8339

مجلة الآداب، مج (٣٤)، ع (٣)، ص ص ٣٩-٥٩، جامعة الملك سعود، الرياض (٢٠٢٢م/ ١٤٤٣هـ)
Journal of Arts, Vol. 34 (3), pp 39-59, © King Saud University, Riyadh (2022 /1443H.)

البيانات الضخمة وأهميتها في العمل الأكاديمي من وجهة نظر الأكاديميين بجامعة طيبة: دراسة استطلاعية

صقر بن مويسان العتيبي

أستاذ علم المعلومات المساعد، قسم المعلومات ومصادر التعلم، كلية الآداب والعلوم الإنسانية، جامعة طيبة، السعودية.

(قدم للنشر في ٢٥/ ٥/ ١٤٤٣هـ، وقبل للنشر في ٢٨/ ٧/ ١٤٤٣هـ)

الكلمات المفتاحية: البيانات الضخمة، الأكاديميين، خصائص البيانات الضخمة، تقنيات البيانات الضخمة.

ملخص البحث: تهدف هذه الدراسة إلى استطلاع أراء الأكاديميين عن البيانات الضخمة وأهميتها، وما يتعلق بها من خصائص وأدوات مستخدمة في تحليل البيانات الضخمة، وقد شارك في الدراسة ٢٥ أكاديميا في كلية الآداب والعلوم الإنسانية في جامعة طيبة. وقد اعتمدت الدراسة الأسلوب الوصفي المسحي باستخدام استبيان استقصائي تم توزيعه عن طريق الإنترنت لجمع البيانات من عينة الدراسة، تم تحليل البيانات باستخدام برنامج التحليل الاحصائي SPSS لقياس المتوسط الحسابي، والانحراف المعياري، والنسب المئوية للإجابة على أسئلة البحث. وقد أظهرت النتائج أن غالبية الأكاديميين على دراية جيدة بالبيانات الضخمة، وأن أهم خصائصها من وجهة نظرهم بالترتيب تنازلياً: الحجم، القيمة، السرعة، التنوع، الصدق والسرعة في الانتشار، كما اتفق غالبية الأكاديميين على أهمية البيانات الضخمة وتأثيرها الإيجابي على أداء عمله، وكذلك أظهرت النتائج أن الأكاديميين لديهم معرفة أو خبرة بالأدوات والتحليلات المختلفة المستخدمة في البيانات الضخمة، وتكمن أهمية هذه الدراسة في أنها واحدة من الدراسات القليلة التي أجريت في المملكة العربية السعودية عن البيانات الضخمة، والأولى عن استخدام البيانات الضخمة في المؤسسات التعليمية من جهة نظر الأكاديميين.

Big data and its importance in academic work from the point of view of academics at Taibah University: a pilot study

Saqar Moisan F Alotaibi
Assistant Professor of Information Science, Department of Information Science, College of Arts and Humanities, Taibah University.

(Received: 25/ 5/1443 H, Accepted for publication 28/ 7/1443 H)

Keywords: Academics, Big Data, Big Data characteristics, Big Data techniques.

Abstract. This study aims to explore the academics' opinions about Big Data, its importance and characteristics, and the tools used in Big Data analysis. A total of 25 academics at the Faculty of Arts and Humanities at Taibah University participated in this study. The study adopted a descriptive research approach using an online survey questionnaire to collect data from the sample. To answer research questions, the data were analyzed using SPSS to measure mean scores, standard deviations, and percentages. The results showed that the majority of participants were well aware of Big Data and its characteristics. According to them, the most important characteristics of Big Data could be arranged in following descending order: size, value, velocity, variety, veracity, and virality. The majority of participants also agreed on the importance of Big Data and its positive impact on their work performance. Further, the results showed that the participants have knowledge or experience of the various tools and analyses used in Big Data. The importance of this study lies in the fact that it is one of few studies that have been conducted in Saudi Arabia on Big Data, besides it is the first on the usage of Big Data in educational institutions from academics' point of view.

المقدمة

توجد البيانات الرقمية الآن في مختلف القطاعات والمؤسسات وأصبحت التكنولوجيا الرقمية تؤدي دوراً رئيسا في الاقتصاد القومي. ورغم أن هذا الموضوع كان ذات يوم لا يَهُم إلا عددا قليلا من خبراء البيانات، فإن البيانات الضخمة أصبحت الآن وثيقة الصلة بصناع السياسات ومتخذي القرارات في مختلف أنحاء المجالات (Manyika et al., 2011). إلى جانب التغيرات التكنولوجية في عصر المعلومات وتكنولوجيا الاتصالات، فقد انتشر علم البيانات الضخمة في معظم المجالات. لقد أصبح مفهوم البيانات الضخمة مصطلحاً عالمياً يُستخدم في مختلف القطاعات الحكومية والتجارية والصناعية والصحية والاكاديمية وغيرها (العتيبي، ٢٠٢٠). إذ بمقدور البيانات الضخمة القائمة على التكنولوجيا وإدارة المعلومات والتحليلات التنبؤية تعزيز كفاءة العمل ومستوى الخدمة (العتيبي،٢٠٢١). وفي ظل التطورات التكنولوجية الحديثة وثورة علم البيانات نجد أن نجاح أي مؤسسة مقترن بالاستفادة من البيانات الضخمة المتاحة لديها (الغبيري وحسن، ٢٠١٩). هذا ولا تقتصر استخدامات البيانات الضخمة على صناع القرار فحسب، بل على المهنيين العاملين أنفسهم.

وقد أثرت التحسينات والتحديات في القرن الحادي والعشرين على مجالات عدة لتضمين تقنيات وتطبيقات تكنولوجيا المعلومات والاتصالات في الأنشطة العادية من أجل تقديم خدمات ناجحة تتسم بالكفاءة (العتيبي،٢٠٢١). إلا أن الضغوط الرامية إلى تنفيذ تغييرات جديدة بعيدة المدى في التعليم العالي تتسبب فيها قوى العولمة الفردية والمشتركة (Daniel, 2017). كما أن انتشار الأشكال الناشئة من بيئات التعليم المعززة للتكنولوجيا (Yuan, Powell & Cetis, 2013) دفع الجامعات للبحث عن طرق عملية للإفادة من بياناتها لمساعدتها على وضع استراتيجيات يمكنها استخدامها للقيام بتغييرات جديدة. ونظراً لوجود البيانات الضخمة بشكل كثيف في العصر الحديث، يحتاج العاملين في المؤسسات التعليمية إلى معرفة العناصر الأساسية للبيانات الضخمة (العتيبي،٢٠٢١) وكيفية تأثيرها في البيئة والعملية التعليمية وعلى العاملين والأكاديميين والطلبة. ولعقود عديدة، وقد مارس التعليم العالي دوماً اتخاذ القرارات القائمة على البيانات، ولكن نادراً ما تم التحقيق في مثل هذه الممارسة بشكل منهجي (Menon, Terkla & Gibbs, 2014; Terkla, Sharkness, Conoscenti & Butler, 2014).

مشكلة الدراسة

تأتي المملكة العربية السعودية في طليعة الدول التي شاركت في بلورة أهداف التنمية المستدامة منذ بدايتها، حيث تضعها في مقدمة أولوياتها بما يتفق مع خصوصيتها وثوابتها (الغبيري وحسن، ٢٠١٩). ولا شك أن لدى المملكة العربية السعودية كما هائلا من البيانات الضخمة في كافة قطاعاتها الحكومية، لذا أصبحت المؤسسات والمنظمات تسعى لتوظيف محللي البيانات من أجل تحليل البيانات المتولدة والمتدفقة والاستفادة منها على أكمل وجه (العتيبي، ٢٠٢٠). كما أن قطاع التعليم العالي يتعرض لضغوط متزايدة لتغيير هيكله التشغيلي والإداري واستيعاب برامج اقتصادية واجتماعية وثقافية جديدة ذات صلة بالمطالب الإقليمية والوطنية والدولية (Daniel, 2017). وبالإضافة إلى ذلك، لا يمكن إنكار أن قطاع التعليم العالي شهد على مدى العقد الماضي نمواً واسعاً في تكنولوجيات التعلم في جميع مجالات التعلم والتعليم تقريبا، مما خلق الإثارة والشواغل بين أعضاء هيئة التدريس.

وعلى الرغم من ذلك، هناك نقص في المواهب التحليلية والإدارية اللازمة للاستفادة إلى أقصى حد من البيانات الضخمة (Manyika et al., 2011). وبالرغم من وجود بعض المقالات عن البيانات الضخمة وأهميتها وفوائدها

وخصائصها وتأثيرها في تعزيز المنافسة بين شركات الأعمال وكيفية الاستفادة منها في عملية صنع القرار، لكن غالبية هذه المقالات نظرية تستعرض الأدب النظري. إلا أن الباحث استطاع إيجاد دراستين إحداهما للغبيري وحسن (٢٠١٩) التي هدفت إلى فهم تأثير البيانات الضخمة على الاتجاه نحو تبني استخدامها في التخطيط لرؤية المملكة العربية السعودية، أما الدراسة الثانية فهي لـ Sreenivasan (٢٠١٧) التي هدفت إلى تحديد الخصائص الأساسية للبيانات الضخمة من وجهة نظر عدد من الممارسين والباحثين في مجال البيانات الضخمة. باستثناء هاتين الدراستين لا تتوافر أي دراسات أخرى على حد علم الباحث عن استخدام البيانات الضخمة في المؤسسات التعليمية وأهمية دورها في تعزيز وتيسير العمل الأكاديمي والبحثي.

أهداف وأسئلة البحث

إن الغرض من الدراسة الحالية هو استطلاع آراء الأكاديميين حول البيانات الضخمة وأهميتها وما يتعلق بها من خصائص وتقنيات ومهارات لذا تتمثل أهداف الدراسة في ما يأتي:

١.اكتشاف تصورات الأكاديميين في جامعة طيبة حول البيانات الضخمة.

٢.تحديد أهم خصائص البيانات الضخمة من وجهة نظر الأكاديميين.

٣.التعرف إلى أهمية البيانات الضخمة من وجهة نظر الأكاديميين.

٤.التعرف إلى الأدوات والتحليلات المستخدمة في البيانات الضخمة.

وبناءً عليه تسعى هذه الدراسة إلى الإجابة على أسئلة البحث الآتية:

١.ما درجة معرفة الأكاديميين في جامعة طيبة بالبيانات الضخمة وما طبيعة استخدامهم لها؟

٢.ما أهم خصائص البيانات الضخمة من وجهة نظر الأكاديميين في جامعة طيبة ؟

٣.ما أهمية البيانات الضخمة من وجهة نظر الأكاديميين في جامعة طيبة؟

٤.ما أدوات وتحليلات البيانات الضخمة التي يستخدمها الأكاديميون في جامعة طيبة؟

أهمية البحث

من الناحية النظرية تكمن أهمية هذه الدراسة في أنها تضيف إلى الإطار النظري في مجال البيانات الضخمة على صعيد مهم وهو التعليم العالي. وعلى وجه الخصوص تسهم هذه الدراسة في معرفة وفهم وجهة نظر شريحة أساسية ألا وهي أعضاء هيئة التدريس العاملين في المجال الأكاديمي الممارسين والباحثين في البيانات الضخمة. أما من الناحية العملية فسوف تساعد هذه الدراسة في فهم تصورات الأكاديمين للبيانات الضخمة وخصائصها، وكذلك أهمية استخدامها في العمل الأكاديمي والبحث العلمي، بالإضافة إلى إدراك القدرة الاستيعابية للبيانات الضخمة والفوائد المدركة لتقنيات وأدوات البيانات الضخمة. مما سيساعد صناع القرار على تبني استخدام البيانات الضخمة في مؤسسات التعليم العالي والاستفادة منها للنهوض بالبيئة التعليمية وكذلك حث الأكاديميين لاستخدامها لتسهيل القيام بأداء مسؤولياتهم الوظيفية.

الإطار النظري:

تعريفات البيانات الضخمةومصادرها:

يُستخدم مصطلح البيانات الضخمة أحياناً كمصطلح مماثل لمفاهيم أخرى مثل استخراج البيانات، وذكاء الأعمال (Su, n.d). تُعرف البيانات الضخمة على أنها كميات ضخمة من البيانات المنظمة/ المهيكلة وغير المنظمة/ غير المهيكلة التي تفيض وتغمر أي مؤسسة، وتزيد كمية البيانات هذه بصورة

كبيرة يومياً (Ajah & Nweke, 2019). تتضمن البيانات الضخمة مجموعات بيانات بأحجام أكبر من قدرة أدوات البرمجيات المستخدمة عموماُ لاستخراج معلومات أو بيانات خلال فترة زمنية محددة (العتيبي، 2021). أما الغبيري وحسن (٢٠١٩) فيُعرفان البيانات الضخمة بأنها كمية هائلة من البيانات المعقدة المتولدة من مصادر بيانية ضخمة الكم، ذات سرعة فائقة وتنوع كبير وحجم ضخم يفوق قدرة البرمجيات والحواسيب الآلية التقليدية على تخزينها ومعالجتها وتوزيعها، وتأخذ أشكالاً متنوعة إذا فُهمت بشكل أعمق واستُخدِمت على نحو أفضل في عملية اتخاذ القرارات. لذا فإن البيانات الضخمة ذات حجم كبير جداً بحيث يصعب مُعالجتها باستخدام التكنولوجيا الحالية والتقليدية لتحقيق الاستفادة منها (البار والمرحبي، ٢٠١٨). في حين أن البيانات الضخمة تنطوي على وجود الكثير من البيانات، إلا أنها لا تشير إلى الحجم وحده، بل أيضاً إلى ميزة الحصول عليها بسرعة عالية، وبصيغ معقدة، ومن مصادر متعددة (Segal, 2019).

قد يكون أحد أهم التعريفات المعروفة للبيانات الضخمة هو تعريف Gartner (٢٠١٢، الفقرة ١) إذ يقول: "البيانات الضخمة هي أصول معلومات تتسم بضخامة كميتها وعلو سرعتها و/أو شدة تنوعها بحيث تتطلب طرق مبتكرة ومجدية التكلفة لمعالجة المعلومات، لتسمح بآراء ثاقبة مُحَسَنة، وإتخاذ القرار، والمعالجة الآلية للعمليات". وفي السياق نفسه فإن De Mauro, Greco & Grimaldi (٢٠١٥، ص ١٠٣) يعدون أن "البيانات الضخمة تمثل المعلومات التي تتميز بكبر كميتها وسرعتها وتنوعها لتتطلب تكنولوجيا وأساليب تحليلية محددة ليتم تحويلها إلى قِيمة". في حين يُعَرف Rouse, Botelho & Bigelow (٢٠١٩، الفقرة ١) البيانات الضخمة بأنها مزيج " من البيانات المهيكلة وشبه المهيكلة وغير المهيكلة التي تجمعها المنظمات... وتستخدمها في مشاريع التعلم الآلي، والنمذجة التنبؤية وتطبيقات التحليلات المتقدمة الأخرى". لذا فالبيانات الضخمة هي تعبير شامل يُستخدم لوصف حجم هائل من البيانات المهيكلة وغير المهيكلة معاً، التي تعد كبيرة وصعبة جداً بحيث يصعب معالجتها باستخدام قاعدة بيانات تقليدية جنباً إلى جنب مع تقنيات البرمجيات.

أما فيما يتعلق بمصادر البيانات الضخمة فيتم جمع المعلومات المشار إليها بالبيانات الضخمة من عدة مصادر مثل قواعد البيانات الجغرافية المكانية (العتيبي، 2021)، وبيانات وسائل التواصل الاجتماعي وبيانات الإنترنت وما إلى ذلك (Rouse et al., 2019)، بل يمكن أن يوفرها الإنسان بشكل مباشر. ويتم خلط هذا النوع من البيانات مع أنواع مختلفة من البيانات والتقنيات المعقدة التي يتم تطبيقها لبناء هذه البيانات (العتيبي، 2021). فعلى سبيل المثال تتكون البيانات المنظمة من معلومات رقمية يتم التقاطها في قواعد البيانات أو مستودعات البيانات أو أوراق العمل؛ أما البيانات شبه المنظمة فتشمل بيانات XML أو سجلات خادم الشبكة أو البيانات المتدفقة من أجهزة الاستشعار،في حين أن البيانات غير المنظمة يتم جمعها من مصادر وسائل التواصل الاجتماعي التي لا تتبع نموذجاً أو تنسيقاً محدداً مسبقاً مثل البريد الإلكتروني، وكتابات نصية وملفات المستندات والصور ومقاطع الفيديو والبيانات الصوتية ورسائل ونقرات على مواقع الإنترنت (البار والمرحبي، ٢٠١٨؛ Rouse et al., 2019 ;Segal, 2019). وهنا يوضح العتيبي (٢٠٢٠) ببساطة أمثلة للبيانات الضخمة منها تغريدات تويتر، ومنشورات وتعليقات فيسبوك، والإعجابات بمنتج أو منشور معين، والرسائل النصية، والفيديوهات، وكذلك مؤشرات الأسهم وأخبار الطقس.

مجالات استخدام البيانات الضخمة

هناك العديد من الطرق التي يمكن بها استخدام البيانات الضخمة لخلق قيمة عبر قطاعات الاقتصاد العالمي. وتشير

الأبحاث إلى أننا أصبحنا على أعتاب موجة هائلة من الإبداع والإنتاجية والنمو، فضلاً عن أنماط جديدة من المنافسة، وكل هذا مدفوع ببيانات ضخمة مع استغلال المستهلكين والشركات والقطاعات الاقتصادية لإمكاناتها (Manyika et al., 2011). وتشكل البيانات دوماً جزءاً من تأثير تكنولوجيا المعلومات والاتصالات إلا أن حجم ونطاق التغيرات التي تجلب البيانات الضخمة بلغت نقطة انقلاب مع تسارع وتقارب سلسلة من الاتجاهات التكنولوجية. وفي واقع الأمر فإن العديد من الشركات الرائدة تستخدم بالفعل البيانات الضخمة لخلق القيمة، ولكن يتعين على شركات أخرى أن تستكشف كيف يمكنها أن تفعل نفس الشيء إذا كان لها أن تنافس. ويجدر الإشارة إلى أن القطاع العام قادر على تعزيز إنتاجيته بشكل كبير من خلال الاستخدام الفعّال للبيانات الضخمة (Manyika et al., 2011).

وتعد هذه البيانات الضخمة قاعدة بيانات كبيرة للغاية يمكن التحقيق فيها بشكل تقديري للكشف عن الاتجاهات والمجموعات والأنماط، لاسيما المرتبطة بسلوك واتصالات الأفراد. ويكشف فحص البيانات الضخمة عن الأنماط والمجموعات والاتجاهات التي يمكن أن تستخدمها أنواع مختلفة من شركات الأعمال وتكنولوجيا الاتصالات والمؤسسات الأكاديمية والمنظمات المختلفة (العتيبي، 2021). بطبيعة الحال فإن هذه البيانات ذات هيكل متباعد طليق إذ تكون غالبًا غير كاملة ولا يمكن الوصول إليها.وتعد البيانات الضخمة هائلة الحجم والسرعة والتنوع في الفوائد بحيث تحتاج إلى أشكال معالجة مبتكرة لتحسين اتخاذ القرارات وتيسير وضع خطط العمليات التشغيلية ورؤى الاكتشاف (العتيبي، 2021). علماً بأن توفر البيانات في أي جهة لا يعني القدرة على استخلاص نتائج فورية، فالبيانات الضخمة تمر بعدة مراحل مـن المعالجة والتنقيـح ليتم تحليلها، ثم استخراج المعلومات المفيدة منها (الغبيري وحسن، ٢٠١٩).

ولا تقتصر استخدامات البيانات الضخمة على مجال تكنولوجيا المعلومات والاتصالات فحسب، بل تمتد إلى مجالات متعددة. فبحسب هاشم، ويعقوب، وأنور، ومختار، وغاني وخان (٢٠١٥) تستخدم البيانات الضخمة في مختلف القطاعات بما في ذلك القطاع الحكومي، القطاع التجاري والمالي، القطاع الصحي والطبي، قطاع الهندسة، قطاع العلوم، وقطاع الصناعة. أماSreenivasan (٢٠١٧) فيشير إلى أنه في السنوات الأخيرة، اكتسبت البيانات الضخمة الاهتمام والتطبيق في مجالات مختلفة مثل الحكومة والرعاية الصحية وقطاع التجزئة والزراعة والبحوث ووسائل التواصل الاجتماعي والاتصالات والبنوك. كذلك يذكر Ajah & Nweke (٢٠١٩) أن البيانات الضخمة تستخدم في الرعاية الصحية، والشبكات، والسفر، وتحليل استهلاك الطاقة، والتطوير التعليمي، والصناعات المالية. وتجدر الإشارة إلى أن مجالات استخدام البيانات الضخمة توسعت أكثر مع ظهور إنترنت الأشياء والذكاء الاصطناعي والحوسبة السحابية، حيث تقوم الشركات بجمع بيانات عن تفضيلات وميول الأفراد.

فعلى سبيل المثال لا الحصر، ذكر كل من البار والمرحبي (٢٠١٨) والتميمي (٢٠١٨) أن استخدام البيانات الضخمة أصبح مهما في القطاع الصحي والطبي إذ يساعد تحليل البيانات الضخمة للمستشفيات والمراكز الطبية والبحثية على التنبؤ بالأمراض والأوبئة واكتشاف مسبباتها مما يسهل اكتشاف العلاج المناسب والسيطرة على الأوبئة ومنعها. أما يوسف (٢٠١٨) فيبين أن للبيانات الضخمة استخدامات في الاقتصاد والمحاسبة تتمثل في مجال القياس المحاسبي حيث إن التقييم والتحليل للبيانات الضخمة يساعد على تطوير معايير المحاسبة واعداد التقارير المالية. كذلك تستخدم البيانات الضخمة في عالم التجارة والأعمال، بدءاً من تجربة العملاء والعمليات والأداء الوظيفي والتحليلات (العتيبي، ٢٠٢٠) مما يساعد الشركات على تطوير المنتجات والخدمات

بشكل أفضل أو ابتكار منتجات جديدة بهدف ضمان مستوى عالٍ من رضا العملاء (Segal, 2019). مما من شأنه أن يساعد في صنع القرارات المستقبلية التي تؤثر في زيادة المبيعات، وتحسين العمل وتقديم خدمات أفضل للعملاء، وتحديد المخاطر وإدارتها (يوسف، ٢٠١٨).

أهمية البيانات الضخمة وفائدتها

إن الاهتمام المتزايد بالبيانات الضخمة يرجع أساساً إلى أنها توفر نظرة قيمة ثاقبة للمنظمات لإصدار القرارات والتوصيات (Sreenivasan, 2017). تكمن أهمية البيانات الضخمة في أنها تستمر بالتولد والتدفق من مصادر متنوعة ومتعددة وبشكل أكبر بكثير مما كانت عليه الحال في السابق. لذلك ومع تزايد كمية وحجم وتنوع البيانات الضخمة التي يتم توليدها وجمعها باستمرار، وجدت الشركات نفسها اليوم أمام خيارين، إما تجاهل هذه البيانات، أو البدء بالتكيف معها تدريجياً لفهمها بهدف التعامل معها والاستفادة منها (حبش، ٢٠١٣). في واقع الأمرفإن البيانات الضخمة في شكلها الخام ليست ذات فائدة كبيرة، ولكن تحليلها يؤدي إلى نتائج قيّمة (Sreenivasan, 2017) وهنا تبرز أهمية وفائدة هذا الكم الهائل من البيانات المتولدة باستمرار.

من فوائد البيانات الضخمة أنها تساعد على حل العديد من المشكلات والأزمات التي تواجه صناع القرار وبالتالي تعينهم على الوصول إلى قرارات سليمة وآراء ثاقبة في وقت زمني قصير بما يلائم طبيعة عمل المنظمات والمؤسسات ونوعية الخدمات أو المنتجات المقدمة، وكذلك الفئة المستهدفة من الزبائن والمستخدمين والعملاء (العتيبي، ٢٠٢٠). وتحديد أفضليات ورغبات فئة عمرية معينة بوساطة تحليل البيانات الخاصة بهم، مما يساعد في تطوير منتجات معينة لهذه الفئة (التميمي، ٢٠١٨). كما يمكن للشركات عن طريق تحليل بيانات العملاء أن تعزز العلاقات معهم بما يحقق ميزة تنافسية. كما تسهم البيانات الضخمة في توفير التكاليف وإنشاء نماذج أعمال جديدة لمنظمات القطاع الخاص والعام (Manyika et al., 2011؛ Ylijoki & Porras, 2016). ومن ثم تحقيق كفاءة عالية في العمليات وزيادة المخرجات التنظيمية (Ajah & Nweke, 2019).

ومن ضمن الفوائد الأخرى للبيانات الضخمة تذكر مؤسسة SAS (٢٠٢٠) أن تحليلها بغض النظر عن مصدرها يساعد متخذي القرار في العثور على إجابات للكشف عن أسباب الفشل والعيوب في وقت قصير، وتخفيض التكلفة والوقت، وإعادة مراجعة حافظات إدارة المخاطر بالكامل في وقت قصير، والكشف عن أي سلوك غير لائق قبل التأثير في المنظمة/ المؤسسة، واتخاذ القرارات الذكية والناجحة، وتطوير المنتجات المحسنة الجديدة. بالإضافة لما سبق فإن البيانات الضخمة تسهم في النمو الاقتصادي من خلال تحفيز الاستثمار وخلق فرص العمل مما يزيد من كفاءة وجودة الخدمات والسلع المنتجة، دعم البحث العلمي والتطوير عن طريق توفير قواعد بيانات تفصيلية للمجالات المبحوثة مما ينتج عنه دقة ومصداقية وكفاءة في نتائج الأبحاث والدراسات، وتشجيع الإبداع والابتكار باستغلال قواعد البيانات الضخمة المتوفرة لتطوير خدمات جديدة ترفع من مستوى أداء المؤسسات أو المنظمات (مركز الاحصاء أبو ظبي، ٢٠١٣).

البيانات الضخمة في التعليم العالي

كان من المحتم أن يؤدي العدد المتزايد من التحديات التي تواجه التعليم العالي إلى قيام المؤسسات التعليمية بجمع المزيد من البيانات واستخدامها (Menon et al., 2014). وقد أشار Wagner & Ice(٢٠١٢)، إلى أن التطورات التكنولوجية كانت بمنزلة عوامل محفزة لنمو تحليل البيانات الضخمة في التعليم العالي. لذا تم اقتراح البيانات الضخمة في السنوات الأخيرة باعتبارها النموذج الأمثل لدراسة ومعالجة التحديات في التعليم العالي (Prinsloo, Archer, Barnes, Chetty & Van Zyl, 2015). وتصف البيانات الضخمة في

التعليم كميات كبيرة من البيانات، سواء كانت على شكل مادي أو رقمي مخزنة في مستودعات متنوعة، تتراوح بين سجلات حسابات ملموسة للمؤسسة التعليمية إلى سجلات اختبارات الطلاب إلى سجلات الخريجين (Sagiroglu & Sinanc, 2013). ويتطلب من مؤسسات التعليم العالي من خلال تواجدها المكثف في العصر الحديث، إلى معرفة أساسيات البيانات الضخمة وكيف يمكن أن تؤثر في العمل الأكاديمي والبحث العلمي والبيئة التعليمية والطلبة.

وعند الحديث عن البيانات الضخمة والتعليم العالي فلا شك أن البيانات الضخمة تؤدي دوراً أكبر بكثير مما ندركه سواء من اختياراتنا للجامعات والكليات أو نمط التعليم أو البيئة التعليمية. وبوصفها علماً جديداً للاستخدام في التعليم العالي، فإن البيانات الضخمة تضم مجالات بحثية مثل استخراج البيانات التعليمية وتحليل التعلم (Daniel, 2017). ويهتم استخراج البيانات التعليمية باستحداث أدوات حسابية لاكتشاف الأنماط في البيانات التعليمية، بينما يركز تحليل التعلم على فهم الطلاب وأدائهم في بيئة تعليمية معينة (Romero, Ventura & García, 2010). وتستخدم الجامعات البيانات الضخمة في عدة طرق مثيرة للاهتمام تفيد في تحقيق الربح أو للعاملين والطلبة على حد سواء (Gutierrez, 2019). وهناك العديد من الطرق التي تجعل التعليم العالي يستخدم البيانات الضخمة لجعل التجربة التعليمية أفضل للطلاب. وتساعد البيانات الضخمة في تحسين الاستجابات لأسلوب التعلم، والاستفادة من التسجيل المستهدف، وزيادة معدلات البقاء (Gutierrez, 2019).

كما أن هناك الكثير من الطرق الأخرى التي تستخدم فيها البيانات الكبيرة في التعليم العالي إذ يمكن للجامعات أن تفيد من البيانات الضخمة لمساعدتها في معالجة إحدى أكبر مشكلاتها ألا وهي الاحتفاظ بالطلاب. إذ من المعروف انخفاض نسب تسجيل الطلاب منذ بعض الوقت، ومعدلات التخرج الفعلية ليست كبيرة بالنسبة للعديد من الجامعات (Gutierrez, 2019). لذا من أجل تحسين احتمالية نجاح الطلاب ودفع المزيد من الرسوم الدراسية، تعمل الجامعات لوقت إضافي لحل هذه المشكلة وتشارك في المعارض التعليمية الدولية للتعريف ببرامجها الأكاديمية وأنشطتها الرياضية والثقافية بهدف جذب المزيد من الطلاب. وهنا يأتي دور البيانات الضخمة حيث أصبح بمقدور الجامعات استهداف طلاب معينين بمجرد أن يعرفوا أن الطالب مهتم بالجامعة أو ببرنامج أكاديمي ما. يمكن للجامعات باستخدام البيانات الضخمة الخوض في درجات الاختبارات والتاريخ الأكاديمي للطالب في المدرسة الثانوية وغيرها من المعلومات لتحديد ما إذا كان من المحتمل أن ينجح (Gutierrez, 2019).

كما بإمكان الجامعات استخدام البيانات الضخمة لترويج برامجها على الإنترنت، وباستخدام هذه المعلومات الإحصائية بمقدورهم تحديد الأماكن التي يمكن فيها تعليم بعض الطلاب على الإنترنت بدلاً من الحرم الجامعي التقليدي (Gutierrez, 2019)، مع العلم بأن التقدم التكنولوجي يسّر للطلاب التواصل مع مدرسيهم بفعالية أكبر من أي وقت مضى من منازلهم. ليس هذا وحسب بل يمكن للجامعات استخدام الخوارزميات لفهم كمية المعلومات التي يجب أن يرسلوها للطلاب بشكل أفضل دون إثقال كاهلهم. بالإضافة لما سبق يمكن للبيانات الضخمة أن تغير طريقة استخدام اختبارات الشخصية إلى جانب المؤشرات الاجتماعية الأخرى ومقاييس الأداء لتحديد بيئات التعلم، وبذلك يستطيع الأكاديميون استخدام هذه المعلومات لتكييف الخطط التعليمية التي ستؤدي إلى تحسين أداء الطلاب وضمان النجاح (Gutierrez, 2019). ويمكن للطلاب أيضا استخدام هذه المعلومات لتحديد أنواع التعليم والمسارات الوظيفية المناسبة لهم وذلك من خلال تقييم

متطلبات سوق العمل ومقارنة تلك المتطلبات بشخصياتهم ورغباتهم التعليمية.

ولوضع تصور للبيانات الضخمة في سياق التعليم العالي، اقترح Daniel & Butson (٢٠١٣) إطاراً نظرياً يصف البيانات الضخمة على أربعة أبعاد وهي: التحليلات المؤسساتية، وتحليلات تكنولوجيا المعلومات، وتحليلات التعلم، والتحليلات الأكاديمية. وتهتم التحليلات المؤسساتية بتحليل البيانات التشغيلية، التي تشمل تحليلات السياسة والتحليلات التعليمية والهيكلية، للمساعدة في اتخاذ قرارات فعالة حول التحسينات اللازمة للمؤسسة التعليمية، حيث يتم استخلاص وتخزين النتائج في مستودعات البيانات (Daniel, 2017). أما تحليلات تكنولوجيا المعلومات فتتعلق ببيانات الاستخدام والأداء الخاصة بالاستخدام المؤسسي لخدمات التكنولوجيا، وتطوير معايير البيانات والأدوات والعمليات، وتهدف إلى دمج البيانات من مجموعة متنوعة من الأنظمة للحصول على نظرة شاملة للاستفادة من الخدمات (Daniel, 2017). وتقوم تحليلات التعلم بقياس وجمع وتحليل البيانات وإعداد التقارير المتعلقة بالمتعلمين أو الطلاب، ويتم إجراؤها على مستوى التدريس والتعلم في المؤسسة التعليمية لأغراض فهم وتحسين عملية وبيئة التعلم ونجاح الطلاب (جونز، ٢٠١٢؛ سيمنز ولونچ، ٢٠١١). أما التحليلات الأكاديمية فتشير إلى تحليلات البرامج الأكاديمية وجميع الأنشطة التي تؤثر في إدارة البرامج وتخصيص الموارد (Tulasi, 2013)، وتوفر معلومات شاملة حول ما يحدث في برنامج محدد وكيفية مواجهة تحديات الأداء إذ أن استخدام المعلومات يساعد الأكاديميين على التعرف إلى الطلاب وبيئات التعلم الخاصة بهم وتوفير التدخلات المستهدفة المطلوبة (Daniel, 2017).

ويمكن للجامعات باستخدام التحليلات التنبؤية تتبع تقدم الطلاب ضمن أهدافهم الرئيسة والوظيفية المحددة وتقييم احتمالية نجاحهم في وقت مبكر من حياتهم المهنية التعليمية (Gutierrez, 2019). كما يمكن للبيانات الضخمة أن تساعد الجامعات في تحديد العوامل غير الأكاديمية التي تجعل الطلاب يرغبون في الانقطاع عن الدراسة. وهذا يساعد الأكاديميين والمستشارين على التواصل ومحاولة تصحيح المسار التعليمي للطلاب وحل المشاكل التي تواجه الطلاب قبل أن تصبح عائقا. ذلك يتضمن في بعض الأحيان توفير خطط تعليمية مخصصة لبعض الطلاب وفي أحيان أخرى يكون التغيير المقترح في التخصص هو الخيار الصحيح والأفضل. ويمكن للمعنيين والخبراء التربويين أن يعملوا على وضع برامج تعمل على تحسين المشاكل المحددة في الحرم الجامعي وجعل البيئة أفضل للجميع (Gutierrez, 2019).

ويمكن ربط الدوافع الرئيسة لتطبيق البيانات الضخمة والتحليلات في التعليم العالي بالضغط المتزايد لإسناد الأدلة على البيانات بدلاً من الحدس أو الخبرة بالإضافة إلى عوامل أخرى مثل: زيادة المساءلة التي يطالب بها الأطراف المعنية، التي بدورها تستلزم جمع أشكال مختلفة من البيانات لإعداد التقارير للوائح الداخلية والخارجية (Daniel, 2017). ويؤدي الاستخدام المتزايد لتقنيات المعلومات المختلفة من قبل الطلاب والأكاديميين والموظفين إلى توليد كمية هائلة من البيانات التي يمكن استخراجها للحصول على معلومات مفيدة. ويعد ظهور أشكال غير منظمة من البيانات الجديدة المرتبطة غالبًا بتقنيات الوسائط الاجتماعية والقيمة التي قد تنجم عن معالجة هذه البيانات (Sagiroglu & Sinanc, 2013) في التعليم العالي سبباً إضافياً لاستخدام البيانات الضخمة للحصول على معلومات مفيدة. إن التبني الناجح لحلول البيانات الضخمة في التعليم العالي يتطلب فهم مجموعة واسعة من البيانات الإدارية والتشغيلية التي تساعد في تقييم الأداء والتقدم وتحديد المسائل المحتملة المتصلة بالبرامج الأكاديمية والبحوث والتعليم والتعلم (Daniel, 2017).

خصائص البيانات الضخمة

يعكس التعريف الرائج للبيانات الضخمة خصائصها الرئيسة وهي الحجم والسرعة والتنوع، إلا أن المزيد من الخصائص تم إضافتها لاحقاً. تقدم كل خاصية من الخصائص ميزة خاصة ومهمة، ففي الحجم هناك تحول في مجموعات البيانات من كبيرة إلى ضخمة، وفي السرعة تحول من المعالجة بالدفعات إلى تدفق البيانات في الوقت الحقيقي، وفي التنوع تحول من البيانات المنظمة إلى البيانات غير المنظمة المتزايدة أو مزيج من الاثنين. ويشير الصدق إلى جودة البيانات، والتباين إلى التفاوت في معدل تدفق البيانات، والقيمة إلى جدوى البيانات في صنع القرار، والصلاحية إلى دقة البيانات. وفي حين أن التقلب هو مدة أهلية البيانات، فالجدوى هي الاختيار الدقيق للسمات، والتصور هو فهم وقراءة البيانات، أما سرعة الانتشار فهي معدل انتشار البيانات.

يستعرض هذا الجزء الخصائص الرئيسة التي ذُكرت في الدراسات السابقة:

•الحجم (Volume): هو حجم البيانات المتدفقة أوالمتولدة من عدة مصادر وهو يحدد قيمة وقدرات البيانات التي سيتم تصنيفها كبيانات ضخمة (Ajah & Nweke, 2019؛ البار والمرحبي، ٢٠١٨). وتأتي البيانات عادةً بكميات ضخمة وبأحجام يتم قياسها بوساطة تيرابايت أو زيتابايت أو إكسابايت (Oracle، ٢٠٢٠). علماً بأن الزيادة في الحجم تعني زيادة في الحاجة إلى تخزين ومعالجة وتحليل البيانات داخل النظام (Riahi & Riahi, 2018)، وتجدر الإشارة إلى أن كميات وأحجام ضخمة من البيانات قد تتسبب في مشكلات إمكانية التوسع وعدم اليقين (Hariri, Fredericks & Bowers, 2019).

•السرعة (Velocity): هي سرعة توليد البيانات والتقاطها وتحليلها لتغطية الطلب عليها (البار والمرحبي، ٢٠١٨؛ Taylor-Sakyi, 2016؛ Rouse et al., 2019)، وهي عنصر حاسم في اتخاذ القرارات (البار والمرحبي، ٢٠١٨)، لذا يجب على سرعة معالجة البيانات أن تلبي سرعة إنتاجها (Hariri et al., 2019). ويتم تحديث البيانات الضخمة بناءً على الوقت الحقيقي، بدلاً من أن يكون يوميا أو أسبوعيا كما يحدث في مستودعات البيانات التقليدية، وهنا تؤدي السرعة دوراً رئيساً، فتقوم أدوات تحليل البيانات الضخمة باستخراجها ومعالجتها وتحليلها لاستخلاص الإجابة (Rouse et al., 2019).

•التنوع (Variety): هو تنوع البيانات المستخرجة سواء المنظمة أوشبه المنظمة أوغير المنظمة المتولدة من مصادر داخلية وخارجية مختلفة (البار والمرحبي، ٢٠١٨؛ Hariri et al., 2019) ويساعد تنوع البيانات المستخرجة الباحثين ومحللي البيانات على اختيار البيانات الأكثر ملاءمة لمجال بحثهم أو اهتماماتهم. ويمكن تخزين الأشكال المختلفة للبيانات في مستودع تخزين للبيانات الخام يُعرف باسم بحيرة البيانات، التي ترتكز على خدمة التخزين الموجهة Hadoop (Rouse et al., 2019). يرتبط التنوع بالاستخدامات المحتملة للبيانات الخام (Riahi & Riahi, 2018)، وأيضاً بتباين البيانات الضخمة.

•التباين (Variability): هو عدم التناسق أو الاختلاف في معدل تدفق البيانات وتعقيد مجموعة البيانات (Sreenivasan, 2017)، التي تشمل كل من البيانات الضخمة الديناميكية والموسمية والمتسلسلة زمنياً والزمانية المكانية وغير الساكنة (صن، ٢٠١٨). كما يعني التباين كيفية توليد البيانات الضخمة من مصادر متغيرة، والتمييز بكفاءة بين البيانات المفيدة المهمة (Kapil, Agrawal & Khan, 2016). ويُقصد به أيضاً المعنى المتغير باستمرار للبيانات وفقاً السياق (McNulty, 2014; Rouse et al., 2019) فيكون للبيانات معاني مختلفة حسب مكان استخدامها، لذا يتم تحليل البيانات الضخمة بتأنٍ ودقة بسبب الطبيعة المتغيرة لمعانيها.

•الصدق (Veracity): هو مصداقية وموثوقية مصادر البيانات التي تتولد منها البيانات (Ajah & Nweke, 2019)، وما شابه من عدم اليقين بسبب عدم الانتظام داخل البيانات (Taylor-Sakyi, 2016). ونظراً لأن الصدق يصف مستوى الدقة وعدم اليقين من البيانات ومصادرها (Riahi & Riahi, 2018)، يجب على محللي البيانات التأكد من مصداقيتها لتحقيق نتائج صحيحة (Rouse et al., 2019) تخدم أهداف المنظمة. بالرغم من مصداقية البيانات الخام الآتية من وسائل التواصل الاجتماعي وصفحات الويب، إلا أن باستطاعة محللي البيانات تحديد وتصحيح أي مشكلات تنجم عن ذلك (Rouse et al., 2019).

•القيمة (Value): هي القيمة والفائدة المُستَخلصة من البيانات (Riahi & Riahi, 2018; Hariri et al., 2019) وتعكس الرؤى المحتملة المستمدة من البيانات (هاشم وآخرون، ٢٠١٥). لا تقتصر القيمة على البيانات المستمدة حصرياً من مصدر البيانات، ولكن من دمج البيانات من عدة مصادر (Su, n.d). بسبب عدم وجود قيمة فعلية لجميع البيانات التي يتم جمعها، يتم إجراء تنظيف للبيانات قبل استخدامها لأغرض العمل (Rouse et al., 2019)، وتقوم الشركات العملاقة بفحص وتحليل كميات ضخمة من البيانات الخام لاستخلاص آراء ثاقبة وقيمة لاتخاذ قرارات أفضل (Hariri et al., 2019).

•الصلاحية (Validity): هي كون البيانات صحيحة ودقيقة للاستخدام المقصود إذ تعد المفتاح لاتخاذ القرارات الصائبة (Khan, Uddin & Gupta, 2014؛ عويس وحسين، ٢٠١٦). وتعد البيانات الضخمة صالحة إذا توافقت مع متطلبات العمل وتم فهمها بشكل صحيح (Sreenivasan, 2017)، وإلا فإن ذلك يؤدي إلى نتائج خاطئة في التحليل وبالتالي أحكام وقرارات خاطئة. ولا تعتمد صلاحية البيانات على اكتمال مجموعة البيانات فقط، بل أيضاً على بيئة العمل والغرض التجاري المفترض أن تخدمه (Khan et al, 2014)، فالهدف منها أن تخبر المستخدم بما تهدف إليه وهل هي صالحة لسياق العمل الحالي أم لا.

•التقلب (Volatility): هي مدة ملائمة وأهلية البيانات الضخمة وفترة تخزينها (Cartledge, 2016). ويعتمد تقلب البيانات الضخمة على الحجم والقيمة والمعالجة وأيضاً المنظمات المختلفة، فمثلاً البيانات الضخمة التي يتم جمعها لأغراض البحث مثل علم الجينوم وعلم الفلك، يتم تخزينها وتحليلها لفترة طويلة (Sreenivasan, 2017). وعند التحدث عن تقلب البيانات الضخمة، يمكننا بسهولة تذكر سياسة الاحتفاظ بالبيانات المنظمة/ المهيكلة التي تستخدمها المؤسسات بشكل يومي في أعمالها (Khan et al, 2014) فبمجرد انتهاء فترة الاحتفاظ، يمكن تدمير أو التخلص من هذه البيانات بسهولة (عويس وحسين، ٢٠١٦).

•الجدوى (Viability): هي عملية الاختيار الدقيق للسمات والعوامل التي من المرجح أن تتنبأ بالنتائج الأكثر أهمية (van Rijmenam, 2013). حيث إن الغرض الأساسي من تخزين وتحليل البيانات الضخمة هو الحصول على نظرة ثاقبة للمساعدة في اتخاذ القرار، يتم تحليل البيانات الضخمة لكشف العلاقات بين المتغيرات المستخدمة في التحليل (Sreenivasan, 2017). يعتقد خبراء البيانات الضخمة أن ٥٪ من سمات البيانات هي المسؤولة عن ٩٥٪ من الفوائد (van Rijmenam, 2013)، لذا لابد من إثبات جدوى البيانات مما يؤدي إلى تقليل مخاطر تكبد تكاليف كبيرة دون توليد فوائد أكبر (Sreenivasan, 2017).

•التصور (Visualization): هو القدرة على قراءة البيانات الضخمة وفهمها مما يجعلها واضحة وسهلة (عويس وحسين، ٢٠١٦). كما يعني كَون البيانات سهلة القراءة والتحليل من الرسوم البيانية المعقدة التي تتضمن العديد من متغيرات البيانات. ويؤدي التصور إلى فهم معنى قيم البيانات المختلفة بشكل أسرع وتحديد العلاقات والأنماط بسهولة (عويس وحسين، ٢٠١٦؛ Sreenivasan, 2017). ويساعد

التصور صانعي القرار –حتى دون خلفية في الإحصاء– على فهم التحليل الاستكشافي الذي يمكن الوصول إليه، مما يجعل عملية تحليل البيانات جزءاً من عملية اتخاذ القرار اليومية (Sreenivasan, 2017).

•سرعة الانتشار (Virality): هي مدى سرعة تفريق المعلومات بين شبكات النظير للنظير وهي تقيس مدى سرعة انتشار البيانات وتقاسمها مع كل عقدة فريدة من نوعها (Wang, 2012)، إذ تعد مقياسا للمعدل الذي تنتشر فيه البيانات في مختلف أنحاء الشبكة (Vorhies, 2014). وتعد سرعة الانتشار عاملاً مهما لنمو العديد من الشركات والأعمال الناشئة على الإنترنت فعند مشاركة المحتوى الخاص بهم (الصوت، الفيديو، التدوينات، التعليقات، التغريدات) وحصوله على إعجاب المستخدمين، تصبح الشركات أو الأعمال مرئية للمزيد من الناس مما يعني الحصول على عملاء جدد بتكلفة أقل (Sreenivasan, 2017).

تحليلات البيانات الضخمة

يشير مفهوم تحليلات البيانات الضخمة إلى استخدام التقنيات التحليلية المتقدمة لتحليل مجموعة البيانات الضخمة باختلاف أنواعها مثل البيانات المنظمة وشبه المنظمة وغير المنظمة المتولدة من مصادرها المختلفة وباختلاف أحجامها أيضاً ابتداء من تيرابايت إلى زيتابايت (أراجيك، د.ت.؛ منصور، ٢٠٢١). وتعتمد آلية عمل تحليل البيانات على قيام الشركات باستخدم منصات التخزين كمحطة أولية لتخزين البيانات الضخمة، ثم نقلها لقاعدة البيانات التحليلية حتى تصبح كاملة وجاهزة للتحليل، ثم يتم تحليلها بوساطة برامج تحليل خاصة عالية الجودة وهي متوفرة بكثرة (أراجيك، د.ت.). وتتميز تحليلات البيانات الضخمة بالتنوع والتطور مما يُمَكِن استخدامها مثل التحليلات التنبؤية (predictive analytics) والتعلم الآلي (machine learning) واستخراج/ التنقيب عن البيانات (data mining) والنمذجة التنبؤية (predictive modeling) وتحليلات النصوص (text analytics) وبرامج التحليل الرياضي والإحصائي. وتكمن مهمة هذه التحليلات في فحص أكبر كمية من البيانات للبحث داخلها واكتشاف أنماطها المختلفة والارتباطات فيما بينها والرؤى الجديدة والعمل على تطوير الرؤى القديمة، مما يؤدي إلى توفير الجهد ويكرس الكفاءة بعكس حلول الذكاء التقليدي (منصور، ٢٠٢١).

ويذكر موقع أراجيك (د.ت.) ومنصور (٢٠٢١) عددا من النقاط فيما يتعلق بأهمية تحليلات البيانات الضخمة فهي تساعد أصحاب المؤسسات على اكتساب الرؤى الجديدة، كما تساعد المنظمات والمؤسسات على تسخير ما تملكه من بيانات وتوظيفها في خدمة الفرص الجديدة المتاحة والتعريف بطريقة استخدامها المثلى. بالإضافة لذلك فإن تقنيات تحليلات البيانات الضخمة المستندة إلى مصادر البيانات مثل السحابة توفر العديد من المزايا فيما يخص التكلفة إذ تعمل على تقليل التكاليف وخاصةً فيما يتعلق بتخزين أكبر كميات ممكنة من البيانات. أما فيما يتعلق بإتخاذ القرارات فبفضل التحليلات التي يتم إجراؤها جنباً إلى جنب مع تحليل المصادر الجديدة للبيانات تساعد هذه التقنيات المحللين والمختصين على اتخاذ القرارات بشكل أسهل وأسرع وأفضل. ليس هذا وحسب بل إن تقنيات تحليلات البيانات الضخمة توفر إنتاج أفضل حيث إن تحليل البيانات يساعد على معرفة احتياجات العملاء وتحديد رضاهم بشكل أوضح، وبالتالي توفير الأفضل سواء في الخدمات أو المنتجات الخاصة بهم مما يؤدي إلى زيادة الإنتاجية.

أدوات البيانات الضخمة

إن تحليلات البيانات الضخمة تسهم في تخزين البيانات وفهمها وتحليلها لاستخراج الأفضل منها لسير العمل المؤسسي، ولضمان نجاح هذه العملية يتم اللجوء إلى برامج أو أدوات تحليل البيانات الضامنة لعملية التحليل. يقول Mazumder (٢٠١٦) إن مشاكل البيانات الضخمة في العصر الحديث كانت مدفوعة بثلاثة تحولات أساسية في

التكنولوجيا والأعمال في العقدين الماضيين وهي: (١) أصبح التخزين الرقمي أكثر فعالية من الورق لتخزين المحتوى وأقل تكلفة، (٢) المعدل غير المسبوق لإنشاء واستهلاك البيانات عبر الويب وإنترنت الأشياء، (٣) الاحتياجات المتزايدة لرصد أنشطة الأعمال على المستويين الجزئي والكلي والتنبؤ بها. وهكذا استلزمت هذه التغييرات في النهاية ظهور العديد من أدوات البيانات الضخمة في مجال إدارة وتحليل البيانات الضخمة. ومن أدوات البيانات الضخمة الأكثر شيوعاً واستخداما Hadoop، HPCC، Apache، Spark، Flink، Storm، Hive، Sqoop، mahout، وغيرها. بالنظر إلى الابتكارات السريعة التي تحدث في أدوات البيانات الضخمة فإن تصنيف أدوات البيانات الضخمة المذكورة أعلاه يستند إلى أن هذه التكنولوجيات يتم نشرها وإدارتها عادة بشكل منفصل لمعالجة مختلف أنواع متطلبات البيانات الضخمة.

علماً بأن أدوات البيانات الضخمة تدعم مختلف أنواع نماذج البيانات لتسهيل الوصول إليها ومعالجتها. بغض النظر عن النماذج، تم تطوير معظم أدوات البيانات الضخمة بناءً على مفهوم تعريف نموذج وقت التشغيل المعروف باسم ' المخطط عند القراءة'(Mazumder, 2016) وذلك يساعد في استيعاب البيانات مقدماً دون اتخاذ قرار بشأن نموذج البيانات أولاً. وبمجرد توفر البيانات في مخزن البيانات، يتم تحديد المخطط (نموذج البيانات) أثناء قراءة البيانات، ويعد هذا النهج مُغيرا رئيسا مقارنة بالنهج المستخدم في تقنيات قواعد البيانات التقليدية حيث يجب تحديد المخطط قبل نقل البيانات إلى النظام الأساسي (Mazumder, 2016). تساعد هذه الفكرة في تعزيز فكرة جلب البيانات أولاً في نظام أساسي مشترك، دون قضاء الكثير من الوقت في التحليل لتحديد مخطط مصادر البيانات الجديدة أو المحدثة (يحدث التحديث كثيرًا بسبب نظام الأعمال المتغير باستمرار). كذلك بمجرد وجود البيانات في مخزن البيانات الضخمة، يتم تنظيف البيانات للحصول على المعلومات ذات الصلة ومعالجتها لتكوين رؤى يمكن استخدامها في النهاية (Mazumder, 2016).

المنهجية

اعتمدت هذه الدراسة الاستطلاعية منهج البحث الكمي باستخدام أسلوب الوصف المسحي. غالباً ما يشار إلى هذا النوع من الأبحاث باستطلاع الرأي أو الاستقصاء، الذي يتكون عادة من أسئلة بسيطة تعطي نتائج رقمية تمكن الباحث من تقديم تفسيرات صحيحة من خلال مقارنة تلك الأرقام (سكيك، ٢٠١٦). ويُعد الأسلوب الوصفي المسحي الأنسب لإجراء هذه الدراسة بسبب طبيعة البحث وأسئلته ومجتمع الدراسة. وتصف الدراسة الاستطلاعية بشكل أساسي خصائص مجتمع الدراسة، حيث يتم دراستها في وقت واحد وخلال مدة زمنية قصيرة نوعا ما (العتيبي وسكيك، ٢٠١٩). استخدم الباحث استبيان استقصائي عبر الويب كطريقة لجمع البيانات من المشاركين في الدراسة.

بيئة الدراسة وعينتها

تم إجراء الدراسة في جامعة طيبة الواقعة في المدينة المنورة في المملكة العربية السعودية. أما بالنسبة لعينة الدراسة فقد شملت ما مجموعه 25 من أعضاء هيئة التدريس العاملين في قسم المعلومات ومصادر التعلم في كلية الآداب والعلوم الإنسانية. أما بالنسبة لاختيار العينة فقد اتبع الباحث نهجاً هادفاً أو ما يُعرف بالعينة القصدية لتطبيق الدراسة كونها دراسة استطلاعية تستهدف أعضاء هيئة التدريس بعينهم.

أداة البحث

بعد اطلاع الباحث على الأدب النظري والدراسات السابقة تم إعداد أداة البحث التي تمثلت في استبيان استقصائي. قام الباحث ببناء استبيان الدراسة الحالية من

استبيانات دراسة الغبيري وحسن (٢٠١٩) ودراسة Sreenivasan (٢٠١٧) مع تعديل بعض الجمل ليناسب الاستبيان الدراسة الحالية، ومن ثم قام مهنيون وأكاديميون ذوو خبرة في البيانات الضخمة بتحكيم الاستبيان بناء على أهداف وأسئلة الدراسة الحالية. بعد إجراء التعديلات اللازمة، تم تصميم الاستبيان باستخدام نماذج جوجل (https://docs.google.com/forms) التي تتيح بناء الاستبيان وتوزيعه إلكترونياً من خلال رابط عبر البريد الإلكتروني، ثم تم إرسال الاستبيان إلكترونياً إلى أعضاء هيئة التدريس مع شرح الغرض من الدراسة ودعوتهم للإجابة على أقسام الاستبيان.

تكون الاستبيان من أربع أقسام إذ شمل القسم الأول المعلومات الشخصية (الجنس، والعمر، والمؤهل العلمي، والوظيفة)، والقسم الثاني أسئلة لكشف معرفة المشاركين بالبيانات الضخمة وتصنيفهم (ممارس، وباحث، ومهتم/ متابع)، وتحديد أهم خصائصها، بينما احتوى القسم الثالث ستة جمل لقياس موافقة المشاركين على أهمية وفوائد البيانات الضخمة في أداء المهام الوظيفية من العمل الأكاديمي والبحث العلمي، أما القسم الرابع فاحتوى خمس جمل لقياس مهارة المشاركين فيما يخص الأدوات والتحليلات المستخدمة في البيانات الضخمة وكفاءات المشاركين. تم تصميم مقياس القسمين الثالث والرابع بناءً على مقياس ليكرت المكون من ٥ نقاط (١= غير موافق بشدة، ٢= غير موافق، ٣= لا ينطبق/ محايد، ٤= موافق، ٥= موافق بشدة).

جمع البيانات والتحليل الإحصائي

تم توزيع الاستبيان عن طريق الإنترنت لجمع البيانات من عينة الدراسة، وذلك في نهاية الفصل الدراسي الأول من العام الأكاديمي ٢٠٢١/٢٠٢٢، ولكون هذه الدراسة استطلاعية فقد استغرق جمع البيانات أسبوعين فقط. بمجرد انتهاء المشاركين من إجابة الاستبيان، تم نقل البيانات مباشرة إلى برنامج التحليل الإحصائي للعلوم الاجتماعية SPSS 26 لإجراء اختبار الموثوقية ولتحليل البيانات. للإجابة على أسئلة البحث، تم استخدام الإحصائيات الوصفية بما في ذلك التكرارات والنسب المئوية والمتوسطات الحسابية والانحرافات المعيارية والرتب لحساب المعلومات الديموغرافية للعينة وعناصر الاستبيان بأقسامه الثلاث الأخرى.

النتائج

التركيبة الديمغرافية

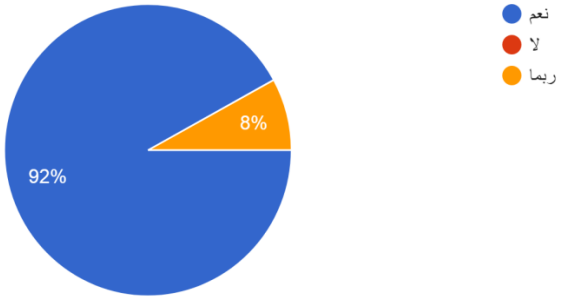
يبين الجدول ١ معلومات المشاركين البالغ عددهم ٢٥ أكاديميا في قسم المعلومات ومصادر التعلم. كما فإن أغلبية المشاركين من الذكور وبلغ عددهم ١٧. أما بالنسبة لأعمار المشاركين فقد تراوحت أعمار الغالبية منهم (١٧ أكاديميا) بين ٣٤ و٤٤ عاماً، يليهم ٥ أكاديميين تراوحت أعمارهم بين ٤٥ و٥٥ عاماً. كما أن الغالبية العظمى من المشاركين يحملون درجة الدكتوراة. أما ما يخص الرتبة الأكاديمية فتعددت ما بين معيد ومحاضر وأستاذ مساعد ووأستاذ مشارك وأستاذ بروفيسور.

جدول ١: معلومات المشاركين الشخصية

العنصر	الفئة	العدد	النسبة ٪
الجنس	ذكر	١٧	٦٨
	أنثى	٨	٣٢
العمر	٢٣-٣٣ عام	١	٤
	٣٤-٤٤ عام	١٧	٦٨
	٤٥-٥٥ عام	٥	٢٠
	٥٦ فما فوق	٢	٨
المؤهل العلمي	ماجستير	٣	١٢
	دكتوراة	٢٢	٨٨
المسمى الوظيفي	معيد	١	٤
	محاضر	٢	٨
	أستاذ مساعد	١٤	٥٦
	أستاذ مشارك	٦	٢٤
	أستاذ/ بروفيسور	٢	٨

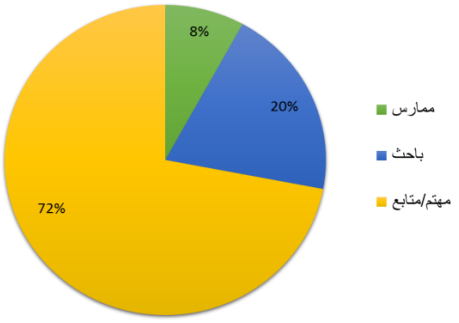
معرفة الأكاديميين بالبيانات الضخمة واستخدامهم لها

ينقسم سؤال البحث الأول [ما درجة معرفة الأكاديميين في جامعة طيبة بالبيانات الضخمة وما طبيعة استخدامهم لها؟] إلى جزئين. للإجابة على الجزء الأول عن درجة معرفة المشاركين بالبيانات الضخمة تم توجيه سؤال مباشر للمشاركين (هل سمعت أو تعرف شيئاً عن البيانات الضخمة؟) وتبين أن ٩٢٪ من المشاركين (٢٣ أكاديميا) على علم بها إذ كانت إجابتهم نعم. وهذا يشير إلى أن الأكاديميين على دراية جيدة وإلمام عالٍ بالبيانات الضخمة وما يتعلق بها. في حين كانت إجابة الاثنين الآخرين ربما كما يبين الشكل ١.



شكل ١: مدى معرفة الأكاديميين بالبيانات الضخمة

أما بخصوص الجزء الثاني من السؤال عن طبيعة استخدامهم للبيانات الضخمة فقد طُلب من الأكاديمين تصنيف أنفسهم إما ممارس أو باحث أو مهتم/متابع لمستجدات الموضوع. كما يتضح من الشكل ٢ أدناه فقد أجاب ٧٢٪ من الأكاديميين (١٨ أكاديميا) بأنهم مهتمين أو متابعين، وذلك بقراءة ما يتم نشره عن البيانات الضخمة، وعد ٢٠٪ منهم (خمسة أكاديمين) بأنهم باحثين يقوموا بكتابة أبحاث أكاديمية ونشرها عن البيانات الضخمة، في حين أجاب ٨٪ (أكاديميَيْن إثنين فقط) بأنهم ممارسين للبيانات الضخمة إذ يستخدمون أدوات وتقنيات البيانات الضخمة في أداء العمل الوظيفي.



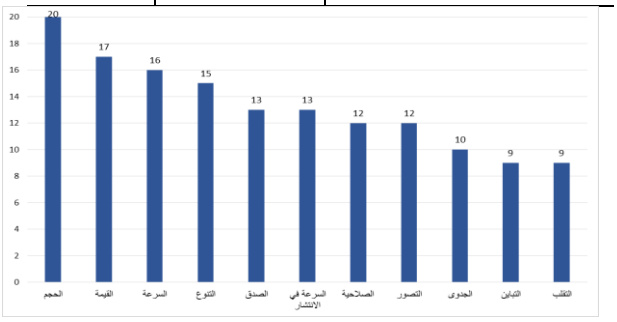
شكل ٢: تصنيف الأكاديميين لأنفسهم

أهم خصائص البيانات الضخمة

للإجابة عن سؤال البحث الثاني [ما أهم خصائص البيانات الضخمة من وجهة نظر الأكاديميين في جامعة طيبة؟] تم وضع ١١ خاصية ذُكِرت في الإطار النظري مع تعريف بسيط وطُلِب من الأكاديميين تحديد أهم الخصائص حسب رأيهم. وبناءً على اختيار المشاركين فإن أهم الخصائص بالترتيب هي الحجم إذ عدها٢٠ مشاركاً أهم خاصية، تليها القيمة كما عدها ١٧ مشاركاً، ثم السرعة كما عدها ١٦ مشاركاً، يليهم التنوع بحسب ١٥ مشاركاً، ويليهم كل من الصدق والسرعة في الانتشار اللذان حازا على ١٣ صوتاً من المشاركين، ومن ثم كل من الصلاحية والتصور اللذان حازا على ١٢ صوتاً من المشاركين. يبين جدول ٢ عدد الأصوات وترتيب الخصائص بحسب الأكاديميين. كذلك يوضح الشكل ٣ ترتيب الخصائص من الأهم إلى الأقل أهمية تنازلياً من وجهة نظر المشاركين.

جدول ٢: ترتيب أهم خصائص البيانات الضخمة من وجهة نظر الأكاديميين

الخاصية	عدد الأصوات من إجمالي ٢٥	الترتيب
الحجم (Volume)	٢٠	١
السرعة (Velocity)	١٦	٣
التنوع (Variety)	١٥	٤
التباين (Variability)	٩	٨
الصدق (Veracity)	١٣	٥
القيمة (Value)	١٧	٢
الصلاحية (Validity)	١٢	٦
التقلب (Volatility)	٩	٨
الجدوى (Viability)	١٠	٧
التصور (Visualization)	١٢	٦
السرعة في الانتشار (Virality)	١٣	٥



شكل ٣: ترتيب الخصائص

أهمية البيانات الضخمة

للإجابة على سؤال البحث الثالث [ما أهمية البيانات الضخمة من وجهة نظر الأكاديميين في جامعة طيبة؟]، تم الطلب من الأكاديميين تحديد درجة موافقتهم على ست جمل تقيس أهمية البيانات الضخمة وفوائدها. تراوحت إجاباتهم بين موافق بشدة ومحايد كما يتضح من الجدول ٣ الذي يعرض عدد إجابات المشاركين بالإضافة إلى المتوسط الحسابي والانحراف المعياري. كما يبين الجدول فإن أدنى متوسط حسابي (١,٣٢) كان للبند الأول (أدرك جيداً أهمية وفوائد البيانات الضخمة في العمل الأكاديمي والبحث العلمي) ما يعكس إدراك الأكاديميين لأهمية البيانات الضخمة في عملهم. أما أعلى متوسط حسابي (١,٩٢) فكان للبند الثاني (تبني استخدام البيانات الضخمة له تأثير إيجابي على أداء عملي) ما يعكس أن الأكاديميين يستخدمون البيانات الضخمة في أداء وظيفتهم.

جدول ٣: أهمية البيانات الضخمة من وجهة نظر الأكاديميين

#	الجملة	موافق بشدة	موافق	محايد	غير موافق	غير موافق بشدة	المتوسط الحسابي	الإنحراف المعياري
١	أدرك جيداً أهمية البيانات الضخمة وفائدتها في العمل الأكاديمي والبحث العلمي.	١٧	٨	٠	٠	٠	1.32	.48
٢	تبني استخدام البيانات الضخمة له تأثير إيجابي على أداء عملي.	٦	١٥	٦	٠	٠	1.92	.64
٣	تؤثر الفوائد المدركة لتحليل البيانات الضخمة تأثيراً إيجابياً في تبني استخدام البيانات الضخمة.	١٠	١٣	٢	٠	٠	1.68	.63
٤	تؤثر الفوائد المدركة للبيانات الضخمة في أداء المهام العملية.	١١	١٢	١	١	٠	1.68	.75
٥	إدراك قيمة البيانات الضخمة له تأثير إيجابي فيالاتجاه نحو تبني استخدام خدمات البيانات الضخمة في العمل الاكاديمي والبحثي.	١٤	٩	٢	٠	٠	1.52	.65
٦	جودة البيانات الضخمة لها تأثير إيجابي قي الفوائد المدركة من تحليل البيانات الضخمة.	١٥	٧	٢	١	٠	1.56	.82

أدوات البيانات الضخمة وتحليلها

للإجابة على سؤال البحث الرابع [ما أدوات وتحليلات البيانات الضخمة التي يستخدمها الأكاديميون في جامعة طيبة؟] تم الطلب من الأكاديميين تحديد درجة موافقتهم على خمس جمل تقيس معرفتهم بأدوات وتحليلات البيانات الضخمة ولغات البرمجة المرتبطة بها، وكذلك كفاءتهم. تنوعت إجاباتهم بين موافق بشدة وغير موافق بشدة كما يتضح من الجدول ٤ الذي يعرض عدد إجابات المشاركين بالإضافة إلى المتوسط الحسابي والانحراف المعياري. كان أدنى متوسط حسابي (٢,١٦) للبند الرابع المتعلق بلغات البرمجة المرتبطة بالبيانات الضخمة ما يعكس قلة خبرتهم بلغات البرمجة. أما أعلى متوسط حسابي (٢,٤٤) فكان للبندين الأول والثاني المتعلقين بدراية المشاركين بالبيانات الضخمة وأدواتها، ما يعكس مدى تمكنهم من البيانات الضخمة.

جدول ٤: أدوات وتحليلات البيانات الضخمة التي يستخدمها الأكاديميين

#	الجملة	موافق بشدة	موافق	لا ينطبق	غير موافق	غير موافق بشدة	المتوسط الحسابي	الإنحراف المعياري
١	أعد نفسي ممارسا أو باحثا مهنيا على دراية جيدة بالبيانات الضخمة.	٥	١٠	٥	٤	١	2.44	1.12
٢	لدي معرفة و/ أو خبرة بواحد أو أكثر من برامج/ أدوات البيانات الضخمة مثل hadoop، HPCC، apache، spark، flink، storm، RapidMiner، pentaho BI، hive، splice، skytree، sqoop، mahout، pig، وغيرها.	٦	٧	٨	٣	١	2.44	1.12
٣	لدي معرفة و/ أو خبرة بواحد أو أكثر من تحليلات البيانات الضخمة مثل التحليلات التنبؤية، النمذجة التنبؤية، والتعلم الآلي، واستخراج البيانات، تحليلات النصوص، وبرامج التحليل الرياضي والإحصائي.	٤	١٤	٥	٢	٠	2.20	.82
٤	لدي خبرة وكفاءة بواحدة أو أكثر من لغات البرمجة المرتبطة بالبيانات الضخمة مثل java،scala، perl،python، C/C++،SQL، pbdR، SAS، MatLab،julia.	٨	٩	٥	٢	١	2.16	1.11
٥	لدي كفاءة في الاستعلام عن تحليل مصادر البيانات المنظمة وشبه المنظمة وغير المنظمة على نطاق واسع.	٤	١٣	٥	١	٢	2.36	1.07

المناقشة

أظهرت النتائج أن الغالبية العظمى من الأكاديميين ملمين بالبيانات الضخمة، كما أظهرت أن طبيعة استخدامهم للبيانات الضخمة تنقسم إما كمهتمين ومتابعين لها أو يقوموا بنشر أبحاث علمية عنها. وتُعد النتائج متوقعة إذ أن البيانات الضخمة أصبحت تُستخدم في مجالات شتى كالتجارة والصناعة والصحة والعلوم والزراعة والتعليم والاتصالات والسياحة والطاقة وغيرها (Ajah & Nweke, 2019، ٢٠١٩؛ العتيبي، ٢٠٢٠؛ Sreenivasan, 2017؛ هاشم وآخرون، ٢٠١٥)، بالإضافة إلى أن المؤسسات والمنظمات في السعودية أصبحت تسعى لتوظيف محللي البيانات الضخمة لتحليلها والاستفادة منها (العتيبي، ٢٠٢٠). لذا من الطبيعي أن يعرف الأكاديميون البيانات الضخمة، وأن يكونوا متابعين لها ويقوموا بنشر أبحاث عنها.

فيما يتعلق بأهم خصائص البيانات الضخمة من وجهة نظر الأكاديمين فهي بالترتيب تنازلياً: الحجم، والقيمة، والسرعة، والتنوع، والصدق والسرعة في الانتشار، والصلاحية والتصور. تتميز نتائج هذه الدراسة بأنها حددت

أهم خصائص البيانات الضخمة بالترتيب بناءً على تقييم الأكاديميين المشاركين على خلاف الدراسات السابقة التي كانت تذكر الخصائص بشكل عام في الإطار النظري، كما تختلف في أن جميع الدراسات السابقة ذكرت أن الخصائص الأساسية تقتصر على الحجم والسرعة والتنوع واختلفوا على درجة أهمية بقية الخصائص.

اتفق غالبية الأكاديميين على أهمية البيانات الضخمة وتأثيرها الإيجابي على أداء واجباتهم الوظيفية سواء العمل الأكاديمي أو البحث العلمي، كما اتفق معظمهم على أن فوائدها المدركة تؤثر إيجابياً في تبني استخدامها وبالتالي على أداء المهام العملية. تتفق هذه النتائج مع ما ذكره العتيبي (٢٠٢٠) ويوسف (٢٠١٨) من أن استخدام البيانات الضخمة يؤثر إيجابياً فيالأداء الوظيفي. كما يتفق مع أهميتها لدعم البحث العلمي مما ينتج عنه دقة ومصداقية في نتائج الأبحاث والدراسات (مركز الاحصاء أبو ظبي، ٢٠١٣)،

فيما يخص أدوات البيانات الضخمة وتقنياتها أظهرت النتائج أن ٥٢٪ من الأكاديميين على معرفة أو خبرة بأدوات البيانات الضخمة المختلفة، ويمكن تفسير هذه النسبة كون الأدوات المذكورة تقع ضمن مجالي علم وهندسة الحاسوب. فيما أكد ٧٢٪ منهم معرفتهم وخبرتهم بالتحليلات المستخدمة في البيانات الضخمة، وتُعد هذه النسبة جيدة جداً كون المشاركين مهتمين وباحثين في مجال البيانات الضخمة، وبالتالي لابد أن يكونوا على علم بالتحليلات المستخدمة. كذلك أجاب ٦٨٪ منهم بأن لديهم خبرة في لغات البرمجة المرتبطة بالبيانات الضخمة وكفاءة في الاستعلام عن مصادر البيانات المنظمة وشبه المنظمة وغير المنظمة وتحليل هذه المصادر ويمكن تفسير هذه النسبة أيضاً لكون هذه اللغات ضمن مجال علم الحاسوب والبرمجة ولعدم خبرتهم كممارسين للبيانات الضخمة، وبالتالي قلة قدرتهم على تحليل مصادرها.

حدود البحث

وجب التنويه إلى أن هذه الدراسة هي دراسة استطلاعية صغيرة، فقد اقتصرت على استطلاع آراء أعضاء هيئة التدريس في قسم المعلومات ومصادر التعلم في جامعة طيبة. كما اقتصرت على فترة زمنية قصيرة وذلك لضيق الوقت بسبب نهاية الفصل الدراسي. لذا من الجيد إجراء دراسات مستقبلية على بقية الأكاديميين العاملين في الجامعة وأن يتم تكرارها في عدة جامعات سواء حكومية أو خاصة، وكذلك استخدام المقابلات الشخصية بجانب الاستبيان للحصول على بيانات نوعية وكمية.

الخاتمة والتوصيات

لا تقتصر االإفادة من البيانات الضخمة على مجال التكنولوجيا، ولكنه يمتد إلى العديد من المجالات مثل الطاقة، والهندسة، والتعليم، والصحة، والصناعة، ومراكز البحث العلمي، والشركات التي تعتمد على البيانات، ومن الواضح أنه سيكون لها آثار مهمة وملموسة على جميع الحقول والمجالات في المستقبل من خلال خلق فرص عمل جديدة ودعم اتخاذ القرار (العتيبي، ٢٠٢٠). كذلك سيصبح الاعتماد على صنع القرار القائم على تحليل البيانات نهجاً مركزياً في العديد من المؤسسات التعليمية والبحثية في التعليم العالي. بل من المرجح أن تعمل البيانات الضخمة والتحليلات على تغيير الطريقة التي يعمل بها التعليم العالي بشكل فعال من خلال استخدام تقنيات مختلفة للحصول على البيانات ومعالجتها وتحليلها وتقييمها واستخدامها لتوليد رؤى قابلة للتنفيذ وصنع قرارات صائبة وثاقبة (Daniel, 2017).

وأخيراً فقد قدمت هذه الدراسة الاستطلاعية الصغيرة نتائج مشجعة فيما يتعلق بأهمية البيانات الضخمة وفائدة تبني استخدامها في العمل الأكاديمي مما يصب في مصلحة الأكاديمين ويساعدهم على التطور وأداء مهامهم الوظيفية

بشكل أفضل وإنتاج الأبحاث العلمية. وعليه فإن قطاع التعليم العالي في المملكة العربية السعودية يجب أن يواكب نظرائه في الدول المتقدمة لتبني استخدام البيانات الضخمة للنمو بالتعليم العالي ووضع الجامعات السعودية في مصاف الجامعات الأفضل والأقوى عالمياً، وبالتالي تصبح السعودية مركزاً تعليمياً إقليمياً ومنافساً قوياً على المستوى العالمي.

وفي النهاية يود الباحث في ضوء نتائج هذه الدراسة التوصية ما يأتي:

١.ضرورة تبني البيانات الضخمة في المؤسسات التعليمية الجامعية لما لها من أهمية للأكاديمين.

٢.الإفادة من نتائج الدراسة في توظيف خدمات البيانات الضخمة في التخطيط والتطوير وصنع القرارات.

٣.تدريب الأكاديمين على استخدام البيانات الضخمة للمشاركة في صنع القرارات الوظيفية وأداء المهام المطلوبة.

المراجع العربية

أراجيك (سنة النشر غير معروفة). ما هو تحليل البيانات الضخمة.

https://www.arageek.com/l/ما-هو-تحليل-البيانات-الضخمة#bfn-footnotes-544769

البار، ع. م.، والمرحبي، خ. ع. (٣ ديسمبر ٢٠١٨). *البيانات الضخمة ومجالات تطبيقها*.

https://www.awforum.org/index.php/en/component/k2/item/190-البيانات-الضخمة-ومجالات-تطبيقها-

التميمي، أ. (٢٠١٨). *أهمية البيانات الضخمة*. متوفر بموقع

https://makkahnewspaper.com/article/1090117/لرأي/ أهمية-البيانات-الضخمة

حبش، م. (٢٤ يوليو ٢٠١٣). *لمحة عن البيانات الضخمة Big Data*.

https://www.tech-wd.com/wd/2013/07/24/what-is-big-data/

العتيبي، ص. م. (٢٠٢٠). البيانات الضخمة: ما هي وما أهميتها وفوائدها وخصائصها المميزة؟ *مجلة دراسات المعلومات، ٢٤*، ٦٩-٨٤.

الغبيري، م. أ.، وحسن، ع. ح. (٢٠١٩). البيانات الضخمة وأثرها في تحقيق رؤية المملكة العربية السعودية ٢٠٣٠ دراسة تطبيقية. *مجلة الاستراتيجية والتنمية*، ٩ (٣)، ٣٢-٥١.

مركز الاحصاء أبو ظبي (٢٠١٣). *مفاهيم عامة حول البيانات الكبيرة: أدلة المنهجية والجودة - دليل رقم (١٣)*.

https://www.scad.gov.ae/MethodologyDocumentLib/13_دليل_البيانات_الكبيرة_النسخة_النهائية.pdf

منصور، إ. (أغسطس ٢٠٢١). *ما هي تحليلات البيانات الكبيرة*. https://almrj3.com/what-is-big-data-analytics/#ref1

يوسف، ج. ع. (٢٠١٨). *مدخل مقترح لتقييم أهمية تطوير المحاسبة في ظل بيئة البيانات الضخمة*.

https://atasu.journals.ekb.eg/article_39108_be7fa5c84add19035db45efdad48e469.pdf

المراجع الأجنبية

Ajah, I. A., & Nweke, H.F. (2019). Big data and business analytics: Trends, platforms, success factors and applications. *Big Data and Cognitive Computing, 3*(32), 1-30. doi:10.3390/bdcc3020032

Alotaibi, S., & Skaik, H. (2019). Univerity digital library usage: A case study from Malaysia. *Modern Directions in Libraries and Information, 26*(51), 179-211.

Alotaibi, S. M. (2021). Big data analysis role in advancing the various activities of digital libraries: Taibah University case study - Saudi Arabia. *IJCSNS International Journal of Computer Science and Network Security, 21*(8), 297-307.

Cartledge, C. (2016). *How many Vs are there in big data?*, 1–4. http://clc-ent.com/TBDE/Docs/vs.pdf

Daniel, B. K. (2017). Big data in higher education: The big picture. In *Big data and learning analytics in higher education* (pp. 19-28). Springer, Cham.

Daniel, B. K., & Butson, R. (2013). Technology enhanced analytics (TEA) in higher education. *Proceedings of the International Conference on*

of the big data. *International Journal of Advanced Computer Science and Applications*, **7**(3), 254-258.

Prinsloo, P., Archer, E., Barnes, G., Chetty, Y., & Van Zyl, D. (2015). Big (ger) data as better data in open distance learning. *The International Review of Research in Open and Distributed Learning, 16* (1).

Riahi, Y., & Riahi, S. (2018). Big data and big data analytics: Concepts, types and technologies. *International Journal of Research and Engineering, 5*(9), 524-528. http://dx.doi.org/10.21276/ijre.2018.5.9.5

Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education, 51* (1), 368–384.

Rouse, M., Botelho, B., & Bigelow, S. J. (2019). *Big data*. https://searchdatamanagement.techtarget.com/definition/big-data

Sagiroglu, S., & Sinanc, D. (2013, May). Big data: A review. In *2013 international conference on collaboration technologies and systems (CTS)* (pp. 42-47). IEEE.

SAS Institute Inc. (2020). *Big data: What it is and why it matters*. https://www.sas.com/en_my/insights/big-data/what-is-big-data.html

Segal, T. (2019). *Big Data*. Available at: https://www.investopedia.com/terms/b/big-data.asp

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review, 46* (5), 30.

Skaik, H. (2016). *Academics' knowledge sharing behaviour in United Arab Emirates*. LAP LAMBERT Academic Publishing, Germany.

Sreenivasan, R. R. (2017). *Characteristics of big data – A Delphi study*. Master's Dissertation, Memorial University of Newfoundland, St. Johan's, NL, Canada.

Su, X. (n.d.). *Introduction to big data: Learning material developed for course IINI3012 big data*. https://www.ntnu.no/iie/fag/big/lessons/lesson2.pdf

Sun, Z. (2018). *10 Bigs: Big data and its ten big characteristics*. https://www.researchgate.net/publication/322592851_10_Bigs_Big_Data_and_Its_Ten_Big_Characteristics

Taylor-Sakyi, K. (2016). *Big data: Understanding big data*. *ArXiv, abs/1601.04602*. https://arxiv.org/ftp/arxiv/papers/1601/1601.04602.pdf

Terkla, D. G., Sharkness, J., Conoscenti, L. M., & Butler, C. (2014). Using data to inform institutional decision-making at Tufts University. In M. E. Menon, D. G. Terkla, & P. Gibbs (Eds.), *Using data to improve higher education* (pp. 39–63). Rotterdam: Sense Publishers. doi: 10.1007/978-94-6209-794-0_4.

Educational Technologies (pp. 89–96.), 29 Novemebr–1 December, 2013, Kuala Lumpur, Malaysia.

De Mauro, A., Greco, M., & Grimaldi, M. (2015). What is big data? A consensual definition and a review of key research topics. *AIP Proceedings of the International Conference on Integrated Information*, 1644, 97-104. doi: 10.1063/1.4907823

Gartner Inc. (2012). *Gartner IT glossary - big data*. https://www.gartner.com/en/information-technology/glossary/big-data

Gutierrez, D. (2019, November). Big data & higher education: How are they connected?. https://insidebigdata.com/2019/11/28/big-data-higher-education-how-are-they-connected/

Hariri, R. H., Fredericks E. M., & Bowers, K. M. (2019). Uncertainty in big data analytics: Survey, opportunities, and challenges. *Journal of Big Data, 6*(44), 1-16. https://doi.org/10.1186/s40537-019-0206-3

Hashem, I. A., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems, 47*, 98–115.

Jones, S. (2012). Technology review: The possibilities of learning analytics to improve learner−centered decision-making. *Community College Enterprise*, *18* (1), 89–92.

Kapil, G., Agrawal, A., & Khan, R. A. (2016). *A study of big data characteristics*. DOI: 10.1109/CESYS.2016.7889917

Khan, M. A., Uddin, M. F., & Gupta, N. (2014). Seven Vs of Big Data understanding Big Data to extract value. *Proceedings of 2014 Zone 1 Conference of the American Society for Engineering Education*. http://www.asee.org/documents/zones/zone1/2014/Professional/PDFs/113.pdf

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

McNulty, E. (2014). *Understanding big data: The seven V's*. http://dataconomy.com/seven-vs-big-data/

Menon, M. E., Terkla, D. G., & Gibbs, P. (2014). *Using data to improve higher education: Research, policy and practice*. London: Springer.

Mazumder S. (2016). Big data tools and platforms. In S. Yu & S. Guo (Eds.), *Big data concepts, theories, and applications* (pp. 29-128). Springer, Cham. https://doi.org/10.1007/978-3-319-27763-9_2

Oracle (2020). *What is Big Data?*. https://www.oracle.com/big-data/what-is-big-data.html

Owais, S. S., & Hussein, N. S. (2016). Extract five categories CPIVW from the 9V's characteristics

Tulasi, B. (2013). Signifi cance of Big Data and analytics in higher education. *International Journal of Computer Applications, 68* (14), 23–25.

van Rijmenam, M. (2013a). *The viability of big data – infographic*. https://datafloq.com/read/viability-of-big-data-infographic/418

Vorhies, B. (2014, October). How many "V"s in big data--the characteristics that define big data. *Data Science Central.* https://www.datasciencecentral.com/profiles/blogs/how-many-v-s-in-big-data-the-characteristics-that-define-big-data

Wagner, E., & Ice, P. (2012). Data changes everything: Delivering on the promise of learning analytics in higher education. *EDUCAUSE Review, 2012* , 33–42.

Wang, R. (2012). Monday"s musings: Beyond the three V"s of big data-viscosity and virality. *A Software Insider's Point of View*. https://www.forbes.com/sites/raywang/2012/02/27/mondays-musings-beyond-the-three-vs-of-big-data-viscosity-and-virality/?sh=6451da3b5fff

Ylijoki, O. and Porras, J. (2016). Perspectives to definition of big data: A mapping study and discussion. *Journal of Innovation Management, 4*(1), 69-91. http://hdl.handle.net/10216/83250

Yuan, L., Powell, S., & Cetis, J. (2013). *MOOCs and open education: Implications for higher education*. http://publications.cetis.org.uk/wp-content/uploads/2013/03/MOOCs-and-Open-Education.pdf\