نورهان سامي فتح الله مرسى. معالجة البيانات الضخمة بالجامعات المصرية باستخدام تطبيقات الحوسبة السحابية: دراسة وصفية تحليلية / نورهان سامي مرسى؛ إشراف بدوية محمد البسيوني، منى فاروق شهوان ـ طنطا: ن. س. فتح الله ، ٢٠٢٢ ـ رسالة ماجستير، كلية الآداب، جامعة طنطا.

عرض
نورهان سامي فتح الله مرسى
مطور مواقع بوحدة البوابة الإلكترونية بجامعة طنطا
norhansamy@unv.tanta.edu.eg

الملخص

تسعى المؤسسات ـــ وخاصة الجامعات ـــ في استثمار المعرفة، والعمل على تحليل البيانات الضخمة، لتحقق لها قيمة مضافة وتساند متخذي القرار بما تنتجه من معلومات ذات أهمية؛ لأن تخزين ومعالجة وتحليل البيانات الضخمة يشكل تحديًا حقيقياً للجامعات والكثير من المؤسسات، نتيجة تنوع وتضخم البيانات بشكل معقد جدا، لذلك فإن التعامل مع البيانات الضخمة **"Big Data"** في المؤسسات يمثل معضلة حقيقية من حيث حجمها المتزايد بشكل سريع، ومن حيث الحاجة إلى مساحات تخزينٍ هائلة تفوق قدرة تلك المؤسسات المالكة لهذه البيانات، وكذلك قلة العائد من هذه البيانات وعدم استغلالها الاستغلال الأمثل، ولهذا فقد أدركت الشركات المتخصصة في إدارة المعلومات وتقنياتها وشركات برمجيات الحاسب الآلي ونظم الأنطولوجيا وإدارة المعرفة، أهمية التعامل بطرق أكثر قدرة على التحكم في هذه البيانات، والإفادة منها بأفضل الطرق، ليتم معالجتها وتحويلها إلى معلومات يستخرج منها معرفة.

وتشكل البيانات الضخمة مع الحوسبة السحابية أكثر ذكاءً وتطوراً، لذلك أصبحت حياتنا تعتمد بشكل كبير على تكنولوجيا الحوسبة السحابية، لإنجاز المهام، حيث أنَّ هناك زيادة سريعة في كمية البيانات الضخمة التي تنتشر في الحوسبة السحابية، وتأتى قوة البيانات الضخمة جنباً إلى جنبٍ مع قوة الحوسبة السحابية، المصحوب بكميات هائلة من البيانات المنشورة على الإنترنت، والتي تحتاج إلى معالجةٍ وتحليلٍ وتخزين.

وتتعدد مصادر هذه البيانات مع ازدياد الوسائل التكنولوجية المنتجة للمعلومات في حياتنا اليومية كالحواسيب، والهواتف الذكية، والمستشعرات الذكية، وآلات الصرف الآلي، والأقمار الصناعية، ومع تحول المؤسسات العامة والخاصة الى استخدام النظم الآلية المتكاملة، وبتعدد هذه المصادر تنوعت الصيغ، ومع استخدام الجميع للتكنولوجيا، والهواتف الذكية، تسارعت عملية إنتاج البيانات بطريقةٍ يصعب معها التعامل مع هذا الكم المتسارع والمتنوع من البيانات، مما أدى إلى ظهور مصطلح البيانات الضخمة.

مشكلة الدراسة

تتمثل مشكلة الدراسة في حدوث تضخم كبير في حجم البيانات والمعلومات المتداولة في الجامعات، ونتيجةً لضخامة هذه البيانات والمعلومات أصبح من الصعب التعامل معها باستخدام قواعد البيانات التقليدية، فيما يعرف بظاهرة البيانات الضخمة "Big Data"، إلى جانب ارتفاع تكاليف تخزين تلك البيانات على خوادم الجامعات، وتكاليف الصيانة، وعمليات الدعم الفني، ومع توجه المنظمات العالمية إلى تخزين البيانات، ومعالجتها، وإدارتها، واستخراج الرؤى منها باستخدام تحليلات البيانات الضخمة، أصبح لزاماً على الجامعات الاستعانة بهذه التقنيات، كتقنيات الحوسبة السحابية، لتحويل هذا الكم الهائل من

البيانات من نقطة ضعف إلى نقطة قوة، ومساعدة مُتَّخذي القرار من اتِّخَاذ القرار السليم، وسط كمٍّ هائل من البيانات.

أهمية الدراسة ومبررات اختيارها:

تتجلى أهمية الدراسة من الاتي:

تستمد الدراسة أهميتها من أهمية الموضوع التي تتناوله، والمتمثلة في محاولة تخزين ومعالجة وتحليل البيانات الضخمة، باستخدام تطبيقات الحوسبة السحابية، وبما أنَّ الحوسبة السحابية هي مستقبل تكنولوجيا المعلومات، ومع انتشار أجهزة الاستشعار، والأجهزة الذكية، وشبكات التواصل الاجتماعي، تضخم حجم البيانات، وأصبح الاعتماد على تطبيقات الحوسبة السحابية وسيلة تساعد على تخزين ومعالجة وتحليل البيانات الضخمة والسيطرة عليها، وصولا إلى دعم اتخاذ القرار السليم.

وبما أن الجامعات تعتبر مؤسسات تسعى إلى حل المشكلات واتخاذ القرار السليم، لذلك تسعى الدراسة لتقدم توصيفاً شاملاً لمدى استخدام تطبيقات البيانات الضخمة السحابية في الجامعات المصرية، مع بيان جدوى تطبيقها، وكيفية التحول من الإدارة التقليدية إلى الاعتماد على البيانات الضخمة في اتخاذ القرار.

أسباب اختيار موضوع الدراسة:

١- قلة الدراسات العربية المعنية بالبيانات الضخمة.

٢- إن البيانات الضخمة بعد معالجتها تتحول إلى معلومات، لذلك فهو مصطلح ارتبط بالمعلومات، التي لا تفصله عن مجال المكتبات والمعلومات، في حين يظل البعض بعيداً عن المصطلح ومدى ارتباطه وأهميته في مجال المكتبات والمعلومات، لذا تهتم هذه الدراسة بإلقاء الضوء على مفهوم البيانات الضخمة، ومجالات الإفادة منها في مجال المكتبات والمعلومات.

٣- الجامعات المصرية كمؤسسات تعليمية تنافسية، تسعى - كغيرها من المؤسسات- للبقاء والنمو، والتكيف مع البيئة الرقمية الجديدة، والتطورات المحيطة بها.

أهداف الدراسة:

تسعى الدراسة إلى تحقيق الهدف الرئيسي **وهو التعرف على دور تطبيقات الحوسبة السحابية في تخزين ومعالجة البيانات الضخمة**، ويتفرع هذا الهدف الرئيسي إلى عدة أهداف فرعية وهي:

١- التعرف على مفهوم البيانات الضخمة ونشأتها وخصائصها ومصادرها وأنواعها وأهميتها واستخداماتها.

٢- التعرف على مفهوم الحوسبة السحابية وأنواعها وأشكالها وخصائصها.

٣- تحديد العلاقة بين البيانات الضخمة والحوسبة السحابية.

٤- حصر وتحليل النظم والبرمجيات المستخدمة في معالجة البيانات الضخمة.

٥- توزيع النظم والبرمجيات محل الدراسة نوعيا وكميا وجغرافيا ولغويا وحسب الإتاحة وقيود الاستخدام.

٦- تحديد الخصائص والإمكانات التي توفرها برمجيات البيانات الضخمة محل الدراسة.

٧- رصد واقع تخزين ومعالجة البيانات الضخمة بالجامعات المصرية محل الدراسة ومدى جهازيتها لاستخدام تطبيقات الحوسبة السحابية.

٨- وضع تصور مقترح لاستخدام تطبيقات البيانات الضخمة السحابية في معالجة البيانات الضخمة بالجامعات المصرية.

حدود الدراسة:

يمكن تحديد نطاق الدراسة من خلال الحدود التالية: -

- **الحدود الموضوعية:** برمجيات البيانات الضخمة السحابية ومدى استخدامها في الجامعات المصرية.
- **الحدود المكانية:** تتناول الدراسة البيانات الضخمة بالجامعات المصرية (جمهورية مصر العربية)
- **الحدود الزمنية:** منذ نوفمبر ٢٠٢٠ حتى مارس ٢٠٢٢.

منهج الدراسة:

تعتمد الدراسة على استخدام:

المنهج الميداني: اعتمدت الدراسة على المنهج الميداني باعتباره أنسب المناهج العلمية لتحقيق ما ترمى إليه الدراسة والذي تم استخدامه في هذه الدراسة لجمع البيانات المتعلقة ببرمجيات البيانات الضخمة السحابية وتحليلها وتفسيرها ، كما تم استخدامه لوصف وتحليل ودراسة الوضع الراهن لمدى استخدام تطبيقات البيانات الضخمة السحابية بالجامعات المصرية، للتعرف على المشكلة عن طريق وصف ظواهرها وخصائصها وطبيعتها ومعرفة أسبابها ومحاولة تفسير العلاقات بين هذه الظاهرة وغيرها من الظواهر وتحليلها ومقارنتها وتفسيرها ثم استخراج النتائج.

أدوات جمع البيانات:

اعتمدت الدراسة على الأدوات التالية لجمع المادة العلمية الخاصة بموضوع البحث:

١- **الاستبيان:** اعتمدت الباحثة على الاستبيان، الذي يُعدُّ الأداة الرئيسية في الدراسات الميدانية بوجه عام، وفى دراسة واقع استخدام تطبيقات البيانات الضخمة السحابية بالجامعات المصرية بوجه خاص، وقد تم تحكيمه وتجريبه مبدئيا على جامعة طنطا، قبل إرساله وتوزيعه على الجامعات المصرية الحكومية محل الدراسة، لتحليل مدى استخدامهم لتطبيقات البيانات الضخمة السحابية

٢- **قائمة مراجعة:** تضم الأسئلة التي على أساسها تم تحليل برمجيات البيانات الضخمة السحابية.

٣- **جلسات الاتصال المباشر بالإنترنت**

مجتمع الدراسة:

يتألف مجتمع الدراسة المستهدف في هذه الدراسة من الجامعات المصرية، والتي تنقسم إلى (الجامعات الحكومية – الجامعات الأهلية – الجامعات الخاصة)، وقد قامت الباحثة باختيار عينة مقصودة وعمدية للدراسة وهى (الجامعات المصرية الحكومية)؛ نظراً لأن الجامعات الأهلية والخاصة حاصلة فقط على الاعتماد من المجلس الأعلى للجامعات، ولا تخضع للإشراف الرئيسي والمتابعة من المجلس الأعلى

للجامعات ووزارة التعليم العالي، حيث يتم الإشراف والمتابعة للجامعات الحكومية من قبل مشروع تطوير نظم وتكنولوجيا المعلومات بالتعليم العالي ICTP، ويكون لها مدير تنفيذي يشرف عليها.

وقد قامت الباحثة بعمل حصر بالجامعات المصرية (الحكومية)، من خلال الموقع الرسمي للمجلس الأعلى للجامعات، وهو كالآتي:

جدول (١) مجتمع الدراسة من الجامعات المصرية وروابطها الالكترونية

م	اسم الجامعة	تاريخ النشأة	الرابط الإلكتروني لها
١	جامعة القاهرة	١٩٠٨	https://cu.edu.eg/ar/Home
٢	جامعة الإسكندرية	١٩٣٨	https://alexu.edu.eg/index.php/ar/
٣	جامعة عين شمس	١٩٥٠	https://www.asu.edu.eg/ar
٤	جامعة أسيوط	١٩٥٧	https://www.aun.edu.eg/main/ar/home
٥	جامعة طنطا	١٩٧٢	https://tanta.edu.eg/
٦	جامعة المنصورة	١٩٧٣	https://www.mans.edu.eg/
٧	جامعة الزقازيق	١٩٧٤	http://www.zu.edu.eg/
٨	جامعة حلوان	١٩٧٥	http://www.helwan.edu.eg/
٩	جامعة المنيا	١٩٧٦	https://www.minia.edu.eg/Minia/
١٠	جامعة المنوفية	١٩٧٦	http://www.menofia.edu.eg/Home/ar
١١	جامعة قناة السويس	١٩٧٦	http://suez.edu.eg/ar/
١٢	جامعة جنوب الوادي	١٩٩٥	https://www.svu.edu.eg/ar/
١٣	جامعة بنى سويف	٢٠٠٥	https://www.bsu.edu.eg//
١٤	جامعة الفيوم	٢٠٠٥	https://www.fayoum.edu.eg/
١٥	جامعة بنها	٢٠٠٥	https://bu.edu.eg/
١٦	جامعة كفر الشيخ	٢٠٠٦	http://www.kfs.edu.eg/
١٧	جامعة سوهاج	٢٠٠٦	https://www.sohag-univ.edu.eg/ar/
١٨	جامعة بورسعيد	٢٠١٠	http://psu.edu.eg/
١٩	جامعة دمنهور	٢٠١٠	http://www.damanhour.edu.eg
٢٠	جامعة أسوان	٢٠١٢	https://aswu.edu.eg/
٢١	جامعة دمياط	٢٠١٢	http://www.du.edu.eg/
٢٢	جامعة السويس	٢٠١٢	http://suezuniv.edu.eg/su/index.php/ar/
٢٣	جامعة مدينة السادات	٢٠١٣	http://usc.edu.eg/
٢٤	جامعة العريش	٢٠١٦	https://www.aru.edu.eg/ar
٢٥	جامعة مطروح	٢٠١٨	http://www.mau.edu.eg/
٢٦	جامعة الوادي الجديد	٢٠١٨	http://nvu.edu.eg/
٢٧	جامعة الأقصر	٢٠١٩	http://www.luxor.edu.eg/

(المجلس الأعلى للجامعات، ٢٠٢٢)

مع العلم أن الباحثة استبعدت أربع جامعات وهي (جامعة العريش – جامعة مطروح – جامعة الوادي الجديد – جامعة الأقصر) للأسباب التالية:

١- حداثة إنشاء تلك الجامعات، حيث تم إنشاء جامعة العريش سنة ٢٠١٦م، وجامعة مطروح سنة ٢٠١٨م، وجامعة الوادي الجديد سنة ٢٠١٨م، وجامعة الأقصر سنة ٢٠١٩، وأيضاً فإنَّ البنية التحتية للأنظمة الآلية لشبكات المعلومات، والبوابة الإلكترونية، ووحدة تطوير النظم لها، لا زالت قيد الإنشاء.

٢- قلة عدد المنتسبين للجامعة من الطلاب وأعضاء هيئة التدريس والعاملين بها، مما يؤدى إلى قلة حجم البيانات، وأنها لا تنطبق عليها خصائص البيانات الضخمة.

كما يشكل مجتمع الدراسة برمجيات البيانات الضخمة المتاحة على شبكة الإنترنت، وذلك من خلال البحث في العديد من محركات البحث وباستخدام المصطلحات باللغة العربية والإنجليزية، ومنها (أدوات البيانات الضخمة – برمجيات البيانات الضخمة – تطبيقات البيانات الضخمة – أدوات معالجة وتحليل البيانات الضخمة - Big data tools - big data applications - big data software - Big Data Analytics Tools) وقد تبين من خلال البحث عن وجود (١٩٠) برنامجا، ولم يتم استبعاد أي من هذه البرامج.

أهم نتائج الدراسة

١- كشفت الدراسة أن الحوسبة السحابية حلاً مناسبًا لمتطلبات تخزين ومعالجة البيانات الضخمة، حيث توفر حلول تخزين قابلة للتطوير وفعالة.

٢- تبين من الدراسة أن الحوسبة السحابية والبيانات الضخمة مكملة لبعضها البعض، فنجد أن الحوسبة السحابية تقنية فعالة توفر حلولا لمشكلة النمو السريع في البيانات الضخمة، في حين أن التخزين التقليدي لا يمكنه تلبية متطلبات التعامل مع البيانات الضخمة، بالإضافة إلى الحاجة إلى تبادل البيانات بين مواقع التخزين الموزعة المختلفة، وهذا ما توفره الحوسبة السحابية.

٣- بينت الدراسة أن علم البيانات يصب كل اهتمامه على **كم البيانات**، في الوقت الذي تشكل **جودة البيانات** أحد الاهتمامات الرئيسية في علم المعلومات.

٤- توصلت الدراسة إلى وجود **١٩٠ برنامجا** من برمجيات البيانات الضخمة، وأن تلك البرمجيات تتنوع وفقا لطبيعة عملها.

٥- أظهرت الدراسة وجود **٦ فئات** لبرمجيات البيانات الضخمة السحابية، لكل فئة خصائصها وسماتها، وهي: (برمجيات التصور المرئي للبيانات - برمجيات معالجة وتحليل البيانات الضخمة - برمجيات قواعد البيانات غير العلائقية NoSQL - برمجيات الذكاء الاصطناعي والتعلم الآلي - برمجيات تخزين البيانات الضخمة - برمجيات معالجة اللغة الطبيعية NLP).

٦- أسفرت الدراسة أن **برمجيات التصور المرئي للبيانات** جاءت في المرتبة الأولى بأعلى نسبة، حيث مثلت **٢٥,٣%** بواقع (٤٨) برنامج من إجمالي برمجيات البيانات الضخمة محل الدراسة، في حين جاءت برمجيات معالجة اللغة الطبيعية في المرتبة الأخيرة بنسبة **١١,١%** بواقع (٢١) برنامج من إجمالي البرمجيات محل الدراسة.

٧- أظهرت الدراسة أن **٧٤%** من الجامعات المصرية محل الدراسة تمتلك بنية تحتية ملائمة من: (الأجهزة – الشبكات – وحدات المعالجة – البرمجيات)، بينما لا تمتلك **٢٦ %** بنية تحتية ملائمة، تمكنها من معالجة البيانات الضخمة.

٨- أظهرت الدراسة أنه **لا توجد** أي استراتيجيات أو خطط واضحة لتطبيق البيانات الضخمة، ضمن الخطط الاستراتيجية للجامعات المصرية محل الدراسة.

٩- أوضحت الدراسة أن أكثر تطبيقات وبرمجيات البيانات الضخمة استخداما في الجامعات **هي منصة Microsoft Azure** ، حيث مثلت أعلى نسبة ٤٢,٦%، يليها في الاستخدام **برنامج Power BI** بنسبة ٢٤%.

١٠-أوضحت الدراسة أن أكثر الجهات المعنية بتقديم الدورات هو **المجلس الأعلى للجامعات**، حيث يمثل أعلى نسبة **٤٦,٧%**، يليه **وزارة التعليم العالي** حيث تمثل نسبة **٢٦,٧%.**

أهم توصيات الدراسة

بناءا على ما توصلت إليه الدراسة من نتائج، فإن الباحثة تقدم مجموعة من التوصيات قامت بتوزيعها على المعنيين بتنفيذها:

توصيات للمتخصصين في مجال المكتبات والمعلومات

١- زيادة الدراسات العربية والاهتمام بإعداد البحوث الأكاديمية حول القضايا المتعلقة بالبيانات الضخمة في المكتبات مثل (تطبيقات البيانات الضخمة في المكتبات – التنقيب عن البيانات وتحليلها ومعالجتها – المكتبات الذكية والرقمية في بيئة البيانات الضخمة – الذكاء التنافسي للمؤسسات وخدمات المعرفة)

٢- الاهتمام بعلم البيانات وبالأخص دراسة البيانات الضخمة داخل أقسام المكتبات والمعلومات بالجامعات من خلال وضع منهج مناسب مع مقدار ما يقدمه مجال المكتبات والمعلومات من أسس نظرية.

٣- على المكتبات أن تأخذ أدورا جديدة تتناسب مع طبيعة البيانات الضخمة وكيفية تحليلها ومعالجتها والتعامل معها بهدف الاستفادة منها وتسهيل الوصول إليها وجعلها في متناول المستفيدين منها، مما يعزز من مستقبل المكتبات على اختلاف أنواعها.

٤- تهيئة البنية التحتية للمكتبات ومراكز المعلومات لتيسير جميع الأنشطة المتعلقة بتطبيق البيانات الضخمة والقيام بتحليلها واستخدامها.

٥- أن تتوافر للمكتبات قوانين مستقلة سواء كانت (سياسة الخصوصية – سياسة أمن المعلومات – سياسة أمن الحوسبة السحابية) وكذلك مواثيق قانونية تنظم أخلاقيات المهنة عند التعامل مع البيانات الضخمة وتحليلها.

٦- ابتكار نماذج فعالة لتطوير دور أخصائيو المعلومات وإعادة تأهيلهم من جديد ليمتلكوا المهارات اللازمة لمواكبة التطورات القادمة في عالم المعرفة.

توصيات للجامعات:

١- تبنى التصور المستقبلي الذي اقترحته الباحثة.

٢- إطلاق الجامعات برامج مميزة جديدة لعلم البيانات وذلك على مستوى درجة الليسانس في المكتبات والمعلومات وكذلك درجة الماجستير والدكتوراه.

٣- طرح مقررات دراسية جديدة في البرامج الأكاديمية لكافة الكليات (على مستوى مرحلة البكالوريوس والليسانس) مثل مقدمة في علم البيانات – معالجة البيانات.

٤- تقديم دورات تدريبية لأعضاء هيئة التدريس والباحثين وطلاب المرحلة الجامعية الأولى على أحدث ما ورد في علوم البيانات.

٥- إنشاء وحدة أو مركز للبيانات الضخمة داخل كل جامعة.

٦- ضرورة الاهتمام بتحديث البنى التحتية للأجهزة والبرمجيات والأنظمة الآلية والشبكات بالجامعات.

٧- ضرورة وضع معايير للبيانات الجامعية لتحديد قيمتها ومدى جودتها.

٨- ضرورة وضع قواعد للتعامل مع البيانات والتأمين والخصوصية كإلزام الجامعات بمنع الاستخدام غير المصرح به للبيانات الشخصية والحساسة.

٩- عمل حصر للعاملين في مجال تكنولوجيا المعلومات IT وخصوصا مجال علم البيانات داخل كل جامعة للقيام بتدريبهم على استخدام تطبيقات وبرمجيات البيانات الضخمة السحابية.