تحليلات البيانات Data Analysis في المكتبات ومراكز المعلومات
مراجعة علمية منهجية

إعداد
علا نبيل القصاص
المدرس المساعد بقسم المكتبات والمعلومات
كلية الآداب ، جامعة المنوفية

الملخص

تهدف هذه الدراسة إلى إستعراض وتحليل أدبيات موضوع تحليلات البيانات في المكتبات ومراكز المعلومات والإلمام بالإتجاهات البحثية المتنوعة للموضوع للتمكن من رسم خريطة كاملة للموضوع متضمنة الموضوعات الرئيسية والفرعية ، للوقوف على نقاط القوة والضعف في كلاً من الإنتاج الفكري العربي والإنجليزي ، والإعتماد على نتائج ذلك في إقتراح إتجاهات بحثية جديدة حول الموضوع . **وذلك بالإعتماد على** المراجعة العلمية المنهجية فهي بحد ذاتها منهج لمعالجة الإنتاج الفكري قائم على التحليل والمقارنة والنقد ، وفي ثنايا تطبيقها تم الاعتماد على كلاً من المنهج الببليوجرافي بأسلوبه المسحي لتجميع الوثائق حول الموضوع ، والأسلوب الوصفي والتحليلي للاطلاع المتعمق لتلك الوثائق والخروج بنتائج فعلية وتوصيات بحثية . **ومن أهم نتائج الدراسة الكمية** أن المنهج الوصفي التحليلي هو الأكثر شيوعاً في معالجة الموضوع وتمثلت نسبة الإعتماد عليه في الإنتاج الفكري العربي بنسبة ٣٨% حيث تحليل واقع المكتبات ومؤسسات المعلومات ورصد نقاط القوة والضعف أو تجارب تمت بالفعل في المكتبات ، اما نسبته في الإنتاج الفكري الإنجليزي ٢٨% لكن تمت معالجته من زاوية مختلفة ألا وهي تحليل تقنيات تحليلات البيانات بشكل مستقل ، إما تحليل لتجربة فعلية من قبل مؤلف المقال .

كما تمثلت نسبة النص الكامل للإنتاج الفكري العربي بنسبة ١٠٠% وأكثر الأعمال نشراً كانت أعمال المؤتمرات بنسبة ٥٤,٥٤%، في حين الإنتاج الفكري الإنجليزي تضمن نص كامل للإنتاج الفكري عامة باستثناء المقالات ، حيث تمثلت نسبة المقالات النص الكامل ٨٨,٨% والمتاح مستخلص فقط بنسبة ١١,١% . **ومن نتائج الدراسة النوعية** أنه تم معالجة الجزء الخاص بدراسة واقع المكتبات ومؤسسات المعلومات ونقاط القوة والضعف حول تطبيق عملية التحليل بشكل وافي ، ولكن وجد قصور في الجزء الخاص بدراسة بخطط المشروعات ، حيث تم تناول الأطر الرئيسية للمشروعات ولكن التطبيق الفعلي للتقنيات والبرمجيات لم تتضح معايير ذلك مثلما تم في الإنتاج الفكري الإنجليزي ، وبالنسبة لإختصاصي المكتبات والمعلومات تم تناول هذا المحور في الإنتاج الفكري الإنجليزي حول أدواره المتنوعة من إيداع ومشاركة نتائج التحليلات ، ولكن في الإنتاج الفكري العربي تناول أهم المهارات الواجب توافرها ورصد المواقع المقدمة لبعض الدورات التدريبية ولاحظت الباحثة عدة فجوات بحثية حول الموضوع وقامت بترجمتها إلى عدة توصيات أو إتجاهات بحثية مستقبلية **ومن هذه التوصيات** ضرورة معالجة الجانب الإقتصادي الناتج عن تحليلات البيانات في المكتبات والتطبيق الفعلي لهذه التقنية الجديدة ، وإجراء الدراسات المقارنة بين تصنيف وتكشيف المحتوى المهيكل وغير المهيكل للمكتبات والدراسات الاستكشافية للخطط الفعلية التدريبية لأمناء المكتبات .

الكلمات المفتاحية : تحليلات البيانات – المكتبات ومؤسسات المعلومات – إختصاصي المكتبات والمعلومات - تقنيات تحليلات البيانات – أدوات تحليلات البيانات .

مقدمة

تعد المراجعة العلمية المنهجية عمل علمي قائم على الاضطلاع على الإنتاج الفكري وإنتقائة وتحليلة ونقدة بشكل تكاملي مقارن ، وذلك لإثراء المعرفة في مجال ما . ومن هنا فقد جاءت هذه المراجعة لتتضمن عرضاً للإنتاج الفكري في موضوع " تحليلات البيانات Data Analysis " ، حيث تم مسح الإنتاج الفكري وتحليلة ووصفة وتقديمة في تجمعات موضوعية متسلسلة لرصد أبعاد الموضوع .

خطوات إعداد المراجعة العلمية

المرحلة الأولى : إختيار الموضوع : تمت مرحلة إختيار الموضوع وفقاً لمحورين :-

المحور الأول : وفقاً لطبيعة المراجعة العلمية

قدم هارولد هارت (Hart , Harold , 1968) رئيس تحرير إحدى دوريات المراجعات الأمريكية مجموعة من المعايير لاختيار موضوع المراجعة العلمية ، ومن بينها معيار تم الاعتماد عليه ألا وهو " النظرة الجديدة " وهذا المعيار يتم الاعتماد عليه في حالة أن الموضوع يتقاطع مع عدد من المجالات المعرفية للخروج بمراجعة نقدية قادرة على إشاعة فكر جديد .

المحور الثاني : وفقاً لتخصص المكتبات والمعلومات

من المعروف أنه كان يتم - سابقاً - تخزين مصادر المعلومات المهيكلة بكافة أشكالها وأنواعها في قواعد بيانات أو مستودعات ، ولكن مع تزايد البيانات بأنواعها المختلفة سواء المنظمة التى تمثل نسبئة ضئيلة أو غير المنظمة التى تمثل نسبة أكبر ، أصبح هناك تحدي متمثل في مقدرة المكتبات ومؤسسات المعلومات فى السيطرة عليها حيث أن تخزين البيانات وتحليلها يتطلب تقنيات حديثة للمساعدة على اتخاذ القرارات .

وبناء على ذلك ، تم اختيار هذا الموضوع للتعرف على الانتاج الفكري العربي و الأجنبي وتحليلة ومعرفة ما انتهى إلية الآخرون ، وطرق معالجتهم للموضوع ، حتى يتثنى نقد ذلك سواء بالايجاب أو بالسلب ، وبالتالي إقتراح موضوعات جديدة لم يسبق البحث فيها .

المرحلة الثانية : مرحلة التخطيط

شملت هذه المرحلة وضع أهداف وحدود ومنهجية الدراسة ، وبحث الإنتاج الفكري العربي والإنجليزي

وتمثلت أهداف الدراسة فيما يلى :-

١- استعراض وتحليل أدبيات موضوع تحليلات البيانات في المكتبات ومراكز المعلومات
٢- التعرف على الإتجاهات البحثية المتنوعة للموضوع
٣- رسم خريطة للموضوع متضمنة الموضوعات الرئيسية والفرعية
٤- إستكشاف نقاط القوة والضعف في كلاً من الإنتاج الفكري العربي والأجنبي
٥- إقتراح عدة إتجاهات بحثية جديدة حول موضوع تحليلات البيانات .

- **تحديد منهج الدراسة**

تم الإعتماد على المراجعة العلمية المنهجية وهي بحد ذاتها منهج لمعالجة الإنتاج الفكري قائم على التحليل والمقارنة والنقد ، وفي ثنايا تطبيقها تم الاعتماد على كلاً من المنهج الببليوجرافي بأسلوبه المسحي

لتجميع الوثائق حول الموضوع ، والأسلوب الوصفي والتحليلي للاطلاع المتعمق لتلك الوثائق والخروج بنتائج فعلية وتوصيات بحثية مستقبلية .

- **حدود الدراسة**

استشهاداً برأي جولي فيرجو (Virgo , Julie , 1971) " المراجعة العلمية المثالية ينبغي أن تزن اسهامات جميع الدول في مجال الاهتمام بموضوعية ، دون القصور في التغطية أو أية تحيزات قومية في الاختيار ، إلا أن المثالية أمر بعيد المنال خاصة لصعوبة بعض الحواجز اللغوية " **وبناء على ذلك تم وضع حدود موضوعية ولغوية وزمنية ونوعية دون وضع حدود جغرافية كالتالي :-**

- **الحدود الموضوعية :** تناولت الدراسة مراجعة أدبيات موضوع " تحليلات البيانات في المكتبات ومراكز المعلومات "
- **الحدود النوعية :** تغطية الرسائل العلمية والمقالات وأعمال المؤتمرات والتقارير
- **الحدود اللغوية :** الإنتاج الفكري العربي والإنجليزي.
- **الحدود الزمنية :** معالجة أدبيات الموضوع في الفترة ما بين ٢٠١٠-٢٠١٩ .

- **بحث الإنتاج الفكرى العربى والإنجليزى**

أولاً : بحث الإنتاج الفكرى العربى

- تم البحث باستخدام مصطلح " تحليل البيانات في المكتبات ومؤسسات المعلومات " وتحديد الفترة الزمنية من ٢٠١٠-٢٠١٩ . وقد تم استخدام **المصطلحات التالية في عملية البحث :**
- **تحليلات البيانات في المكتبات ومراكز المعلومات**
- **تقنيات تحليلات البيانات**
- **أدوات تحليلات البيانات**

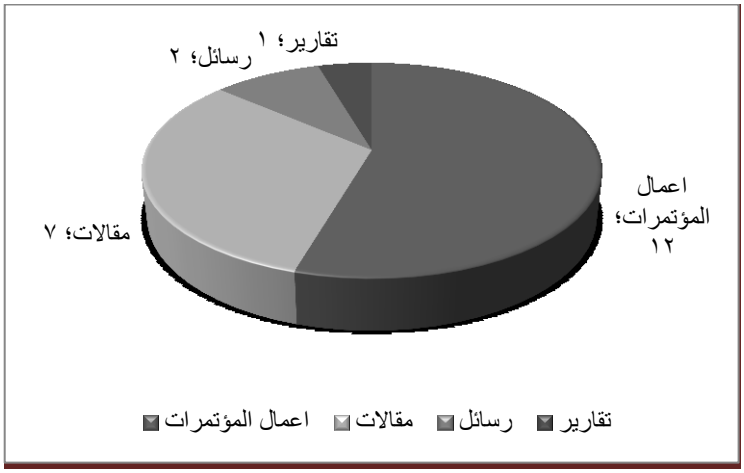
دور إختصاصي المكتبات والمعلومات في تحليلات البيانات

ويوضح الجدول التالى النتائج المسترجعة :-

المصدر	عدد التسجيلات
قاعدة بيانات الهادي	٥
دار المنظومة	٦٢٧٨
Google scholar	١١٠٠٠

بعد تنقية النتائج واستبعاد المكرر منها، واستبعاد التسجيلات التى ليس لها علاقه بالموضوع بشكل مباشر ، واستبعاد المصادر التى يتاح المستخلص فقط ولم يتوافر فيها النص الكامل ، ظهرت النتائج كما تم توضيحها بالجدول والشكل التالى :

نوع المصدر	العدد	النسبة المئوية
أعمال مؤتمرات	١٢	%٥٤,٥٤
مقالات	٧	%٣١,٨١
رسائل	٢	%٩,١١
تقارير	١	%٤,٥٤



إتجهت الباحثة بعد ذلك إلى تحليل المقالات لرصد المناهج المتبعة فى معالجة الموضوع ، وكذلك لحصر أماكن النشر وللوقوف على أكثر المؤلفين إنتاجية . وذلك يمكن إيضاحة كالتالي :-

المنهج المتبع

تنوعت المناهج المتبعة في الإنتاج الفكري العربي بنسب مختلفة ، وكانت النتيجة كالتالي :-

المنهج	عدد المصادر	النسبة المئوية
وصفي تحليلي	٨	%٣٨
وصفي	٥	%٢٣,٨
دراسة حالة	٣	%١٤,٢
تقييمي	١	%٤,٧
وصفي كمي	١	%٤,٧
مسحي ميداني	١	%٤,٧
نوعي	١	%٤,٧
مقارن	١	%٤,٧
الإجمالي	٢١	%١٠٠

ويوضح الجدول السابق المناهج المستخدمة فى الإنتاج الفكرى محل الدراسة ، وبصرف النظر عن الاختلاف فى تسمية المناهج لدى المتخصصين فإن المنهج الوصفى كان أكثر استخداما فى هذا الموضوع فى الإنتاج الفكرى العربى ، وذلك بتسمياته المختلفة (الوصفى التحليلى ـ الوصفى ـ الوصفى الكمى ـ المسحى الميدانى) ويرجع سبب ذلك إلى أن معظم الدراسات عالجت الجزء التحليلي بأحد الجانبين ، إما تحليل للواقع الخاص بالمكتبات ومؤسسات المعلومات ورصد نقاط القوة ونقاط الضعف ، وإما تحليل لوضع المكتبات المطبقة لتلك التحليلات .

مصادر النشر بالإنتاج الفكرى العربى

اقتصرت مصادر النشر على المؤتمر الرابع والعشرون والذى أقامته جمعية المكتبات المتخصصة فرع الخليج عام ٢٠١٨ والخاص بالبيانات الضخمة وآفاق استثمارها والذى أفرز ١٣ بحثا ، وجاءت ابحاث الدوريات العلمية لتفرز ٧ بحوث ، ثم جاءت رسالتين فى جامعتي سوهاج والمنوفية ، والجدول التالى يوضح تفاصيل ذلك :

مصادر النشر	عدد المصادر التي تم نشرها
المؤتمر الرابع والعشرون : البيانات الضخمة وآفاق استثمارها	١٣
مجلة مكتبة الملك فهد الوطنية	٢
المجلة الدولية في مجال الحاسبات والمعلومات جامعة المنوفية IJCI	٢
مجلة المكتبات والمعلومات العربية	١
المجلة العلمية لقطاع كليات التجارة جامعة المنوفية	١
قسم المكتبات والمعلومات جامعة سوهاج	١
قسم المكتبات والمعلومات جامعة المنوفية	١
الإجمالي	٢١

بالنسبة للمؤلفين بالإنتاج الفكرى العربى

تم تطبيق قانون لوتكا ، وكانت النتيجة بأن **" علي بن ذيب الأكذبي "** هو الأكثر نشاطاً في معالجة الموضوع في الفترة الزمنية المحددة (٢٠١٠-٢٠١٩) ، حيث نشر بحثين أحدهما في مؤتمر جمعية المكتبات المتخصصة فرع الخليج والآخر في مجلة الملك فهد الوطنية ، فيما عدا ذلك تم نشر عمل واحد لكل مؤلف .

ثانياً : بحث الإنتاج الفكرى الإنجليزي

تم تحديد مصطلح البحث " Data Analysis in Libraries" في الفترة ما بين ٢٠١٠-٢٠١٩. وقد تم استخدام **المصطلحات التالية في عملية البحث :**

- **Data Analysis in libraries**
- **Data analysis techniques**
- **Data analysis tools**
- **The role of librarians in data analysis**

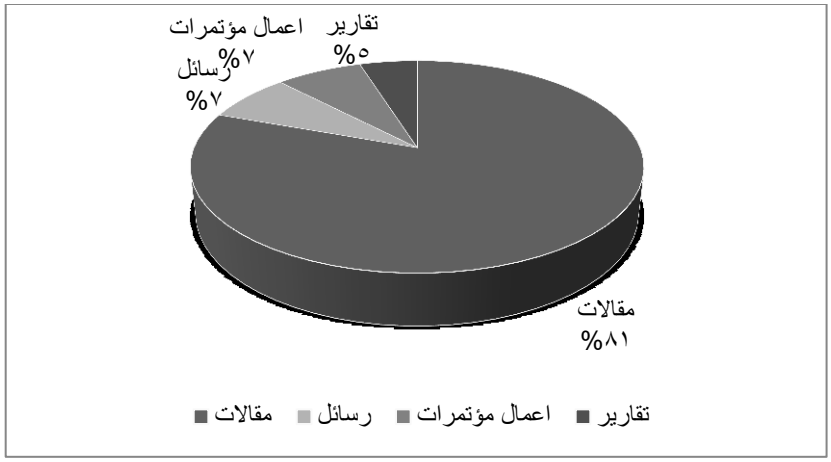
ويوضح الجدول التالى النتائج المسترجعة :-

Source	Result
EBSCO	1,624,916
EMERALD	3100
PROQUEST	20
GOOGLE SCHOLAR	17,800

بعد تنقية النتائج واستبعاد المكرر منها، واستبعاد التسجيلات التى ليس لها علاقه بالموضوع بشكل مباشر ، واستبعاد معظم المصادر التى يتاح المستخلص فقط ولم يتوافر فيها النص الكامل ، حيث تم

الاحتفاظ بنسبة تقدر ١١,١ % من المقالات التى توافرت فيها مستخلصات فقط وذلك لأهمية هذه المقالات وريادتها فى الموضوع وعدم تمكن الباحثة من الحصول على النص الكامل ، وظهرت النتائج كما تم توضيحها بالجدول والشكل التالى :

نوع المصدر	النتائج	النسبة المئوية
مقالات	٣٣	%٨٠,٥
رسائل	٣	%٧,٣
أعمال مؤتمرات	٣	%٧,٣
تقارير	٢	%٤,٩
الإجمالي	٤١	%١٠٠



إتجهت الباحثة كما سبق في الإنتاج الفكري العربي برصد المناهج العلمية المتبعة ، ورصد مصادر النشر في الموضوع والوقوف على أكثر المؤلفين إنتاجية كالتالي :-

المنهج المتبع

المنهج	عدد المصادر	النسبة المئوية
وصفي تحليلي	١٠	%٢٨
تحليلي استكشافي	٨	%٢٢
وصفي	٧	%٢٠
استقصائي ميداني	٣	%٨
دراسة حالة	٣	%٨
تحليل لاحق	٢	%٦
نظري استكشافي	١	%٢
تاريخية	١	%٢
لم تذكر المنهج	٦	
الإجمالي	٤١	% ١٠٠

وبالمقارنة بالمناهج المتبعة في الإنتاج الفكري العربي ، لاحظت الباحثة أن أكثر المناهج استخداما أيضاً هو المنهج الوصفي التحليلي ، ولكن الاختلاف كان فى اسلوب المعالجة ، حيث تم معالجتة من وجهة نظر مختلفة عن الإنتاج الفكري العربي ، حيث " تحليل تقنيات لتحليلات البيانات بشكل مستقل ، إما تحليل لتجارب طبقت بالفعل " أما باقي المصادر لم يذكر بها المنهج المتبع

مصادر النشر في الإنتاج الفكري الإنجليزي

مصدر النشر	عدد المصادر	مصدر النشر	عدد المصادر
Computer standards & interface	١	Procedia computer science	١
Journal of library Administration	٢	Computer research repository	١
Acm sigmetrics performance evaluation review	١	Journal of academic librarianship	٢
IEEE Journal of biomedical and health	١	IEEE signal processing magazine	١
IEEE international computing	٤	Information Systems	١
MC Kinsey global institute	١	IEEE communications survey	١
Journal of super computing	١	International conference on measurement and modeling of computer science	١
IEEE transactions on visualizations and computer graphics	١	5th biennial conference on innovation data systems research	١
International conference on big data and kwnoledge discovery	١	4th international symposium on emerging trends and technologies in libraries	١
British journal of educational technology	١	Journal of documentation	١
School of strategic leadership	١	Massachusetts institute of technology	١
Umi desertation publishing	١		

لاحظت الباحثة خلال الجدول السابق تفوق كلاً من **Journal of library administration – Journal of academic librarianship – IEEE international computing** ، وذلك يدل على إهتمام الدوريات المتخصصة بموضوع تحليلات البيانات وسبب نشرها في دورية إدارة المكتبات يرجع إلى أن تحليلات البيانات أثرت وبشكل واضح على الإدارة سواء إدارة المكتبة ككل أو إدارة البيانات في كل قسم على حدة ، وبالنسبة لدورية القيادة الأكاديمية **ترجع الباحثة السبب إلى** أن تطبيق تحليلات

البيانات يتطلب قيادة فعالة تساهم بإدارة متميزة لمشروع التحليلات ، أما بالنسبة لأكثر الدوريات نشراً وهي IEEE يعد أمراً طبيعياً حيث أن موضوع تحليلات البيانات ظهر بداية بتخصص علم الحاسوب والبرمجيات والتقنيات ومازالت هذه الدراسات حيث نشر أساسيات التعامل مع التحليلات ومن ثم توظيفها في الدوريات الآخرى كل حسب تخصصة .

وبالنسبة **للمؤلفين** في الإنتاج الفكري الإنجليزي ، بعد تطبيق قانون لوتكا ، كانت النتيجة عدم تفوق مؤلف عن باقى المؤلفين ، حيث أن كل مصدر منفرد بمؤلف واحد أو عدة مؤلفين مختلفين عن باقي المصادر .

وجاء التوزيع الزمني والنوعي للإنتاج الفكري العربي والإنجليزي كما يتضح في الجداول التالية :-

أولاً :- الإنتاج الفكرى العربى

التوزيع النوعي التوزيع الزمني	مقالات	أعمال مؤتمرات	رسائل علمية	تقارير	المجموع	النسبة المئوية
٢٠١٢	١	-	-	-	١	%٤,٥٤
٢٠١٤	١	-	-	١	١	%٤,٥٤
٢٠١٧	٢	-	-	-	٢	%٩,١١
٢٠١٨	١	١٢	١	-	١٤	%٦٣,٦٣
٢٠١٩	٢	-	١	-	٣	%١٣,٦٣
المجموع	٧	١٢	٢	١	٢٢	
النسبة المئوية	%٣١,٨١	%٥٤,٥٤	%٩,١١	%٤,٥٤		%١٠٠

لوحظ أن سنة ٢٠١٨ هي سنة ازدهار الإنتاج الفكري باللغة العربية في مجال تحليل البيانات في تخصص المكتبات والمعلومات ، يرجع سبب ذلك إلى إنعقاد مؤتمر جمعية المكتبات المتخصصة فرع الخليج العربي " المؤتمر الرابع والعشرون : البيانات الضخمة وآفاق إستثمارها : الطريق نحو التكامل المعرفي " . وذلك بالطبع نتج عنه تفوق أعمال المؤتمرات عن باقي الإنتاج الفكري العربي ، كما أن الباحثة إستبعدت النتائج التي توصلت إليها خلال عام ٢٠١٦،٢٠١٥،٢٠١٣،٢٠١١،٢٠١٠ نظراً لعدم تعمق محتواها بعد فحصة في موضوع تحليلات البيانات . كما لاحظت الباحثة أن سنة ٢٠١٢ ،٢٠١٤ تعد أقل إنتاج فكري بنسبة **٤,٥٤%** ، وذلك يرجع إلى بداية ظهور الموضوع في الإنتاج الفكري العربي وخصوصاً أن مقاله من الإثنين متعلقه بشكل كامل بالجزء التقني في الموضوع وما زال في تطور والدليل على ذلك أنه يوجد مقالتين في ٢٠١٩ أيضاً في الجزء التقني ، فيما عدا ذلك يصبوا في التخصص بشكل مباشر .

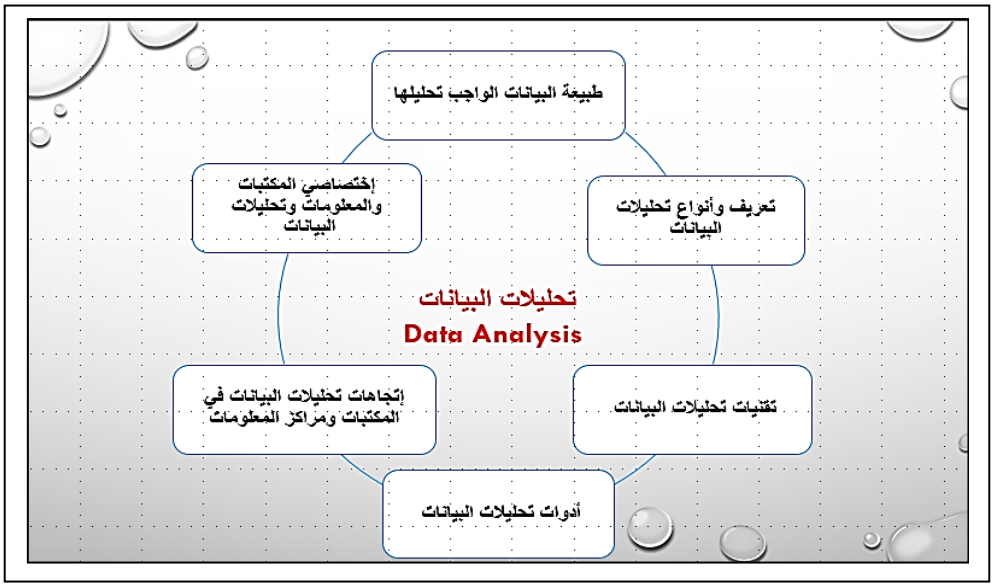
ثانياً :- الإنتاج الفكري الإنجليزي

التوزيع النوعي / التوزيع الزمني	مقالات	أعمال مؤتمرات	رسائل علمية	تقارير	المجموع	النسبة المئوية
٢٠١٠	-	١	-	-	١	%٢,٤
٢٠١١	٧	١	-	١	٩	%٢١,٦
٢٠١٢	٩	-	١	-	١٠	%٢٤,٤
٢٠١٣	٣	-	١	-	٤	%٩,٨
٢٠١٤	٦	-	١	-	٧	%١٧,١
٢٠١٥	٢	-	-	١	٣	%٧,٣
٢٠١٦	٣	-	-	-	٣	%٧,٣
٢٠١٧	٢	-	-	-	٢	%٤,٩
٢٠١٨		-	-	-	-	-
٢٠١٩	١	-	-	-	١	%٢,٤
الإجمالي	٣٣	٢	٣	٢	٤١	
النسبة المئوية	%٨٠,٥	%٧,٣	%٧,٣	%٤,٩		%١٠٠

ولوحظ أن سنة ٢٠١٢ هي سنة إزدهار الإنتاج الفكري باللغة الإنجليزية في مجال تحليل البيانات في المكتبات والمعلومات ، وبعد الإطلاع على المصادر المنشورة خلال هذا العام وجد أن معظمها مقالات وهذه المقالات تتضمن ٨ في التخصصات التقنية وواحدة فقط في تخصص المكتبات والمعلومات ، ويمكن أن ترجع الباحثة السبب في ذلك إلى أن بداية ظهور التقنيات الحديثة في تخصص المكتبات والمعلومات جاء بعد الاطلاع عليها في الدوريات المتخصصه فى المجالات التقنية وتطبيقها في المكتبات ومراكز المعلومات ثم الكتابة عنها فى دوريات المكتبات والمعلومات وذلك يعد أمراً طبيعياً . وبالنسبة لنوعية الإنتاج الفكري الإنجليزي لوحظ تفوق المقالات بنسبة ٨٠,٥% .

ثالثاً : مرحلة الإطلاع والتحليل

قامت الباحثة في هذه المرحلة بالإطلاع على الإنتاج الفكري وتحليل الوثائق تحليل نقدي مقارن ، لتقسيم الوثائق إلى فئات تبعاً لموضوعها وعلاقاتها " رسم خريطة للموضوع "



شكل رقم (١) خريطة مبدئية لمحاور موضوع تحليلات البيانات

قامت الباحثة بعد رسم هذه الخريطة المبدئية بتسجيل مفردات الإنتاج الفكري وعلاقة تقسيماتها الفرعية ببعضها البعض ، ثم كتابة وصياغة المراجعة كما يلي :-

مصطلح تحليل البيانات :

لابد قبل التطرق إلى دراسة موضوع ما ، تعريف المصطلح الأساسي وهو مكون من شقين " تحليل " و " البيانات " ، **والتحليل** لغوياً عرفه معجم الوسيط (مجمع اللغة العربية ، ٢٠١١) بأنه تحليل الجملة : بيان أجزائها ووظيفة كل منها ، والتحليل النفساني فرع من علم النفس الحديث يبحث في العقل الباطن وما فيه من عقد ورغبات تمهيداً لعلاجها . و البيان : الحجة والمنطق الفصيح والكلام يكشف عن حقيقة حال ، أو تحمل في طياته بلاغا و علم يعرف به إيراد المعنى الواحد بطرق مختلفة من تشبية ، ومجاز ، وكناية .

أما كلمة " بيانات " فهى جمع بيان ويقصد بها المعلومات التفصيلية حول شخص أو شيء ما يمكن من خلاله الاستدلال عليه .

المعنى الإصطلاحي لمصطلح تحليل البيانات :

ولم تختلف تعريفات (تحليل) عما جاء فى التعريف اللغوى ، أما مصطلح البيانات فعرفه قاموس الشامي بأنه تمثيل للحقائق والأرقام أو التعليمات المقدمة في شكل مناسب لفهمها وتفسيرها بحيث يناسب عمليات الاتصال أو المعالجة بواسطة الإنسان أو الكمبيوتر .

كما عرف قاموس Business Dictionary مصطلح " تحليلات البيانات " بشكل متصل بأنه عملية تقييم البيانات باستخدام المنطق التحليلي لفحص كل عنصر من مكونات البيانات ، حيث يتم جمع البيانات من مصادر مختلفه ، ومراجعتها ، ومن ثم تحليلها لتشكل نوعا من الاستنتاج ،و هناك مجموعة متنوعة من طرق تحليل البيانات المحددة، يتضمن بعضها استخراج البياناتdata mining، وتحليل النصوص text analytics، وذكاء الأعمالbusiness intelligence، ومرئيات البيانات data visualizations .

وبعد الاطلاع على أدبيات الموضوع باللغة العربية والانجليزية ، وجدت الباحثة أنه لم يتم التطرق إلى تعريف تحليلات البيانات بشكل واضح ، لكن تم تعريف أنواع هذه التحليلات كل حسب الغرض منه ، فقسمها (أحمد خيري ، ٢٠١٨) إلى أربع أنواع وهي :-

- **تحليلات وصفية Descriptive Analytics** : يعد أبسط أنواع التحليلات ويعمل على تحويل الكميات الضخمة من البيانات المتشابكة والمعقدة إلى بيانات سهلة الفهم وذات مغزى ويفيد التمثيل المرئي للبيانات في هذا النوع حيث يسهل عملية الوصف ومن أشهر الأعمال التي تستخدم فيها هذه التحليلات خرائط الظواهر الطبيعية مثل الزلازل والبراكين .

- **تحليلات تشخيصية Diagnostic Analytics** : ينظر هذا النوع في الأحداث الماضية لتحديد ما الذي حدث ؟ ولماذا حدث على هذا النحو ؟ بمعنى آخر تكشف لنا الجذور والاسباب الأساسية التي تسببت في وجود حدث ما ، وغالبا يستخدم هذا النوع في إعتماد الاستراتجيات الثابتة للمنظمة من حيث الميزانية والأرباح وتلافي المخاطر وتغيرات الأداء السلبي وتخفيض الأرباح وغيرها .

- **تحليلات تنبؤية Predictive Analytics** : مجال للتحليل الاحصائي للبيانات هدفه استخلاص معلومات حول التغيرات السلوكية المستقبلية ، بحيث تقوم بتحديد السيناريوهات المستقبلية التي يمكن أن تحدث اعتماداً على ما سبق من تحليلات وصفية وتشخيصية ، ويتأسس التحليل التنبؤي على فهم العلاقة بين المتغيرات التي تتسبب في الأحداث واستعمال هذه العلاقات من أجل توقع المستقبل حيث يقوم بتحديد أنماط البيانات السابقة ويقدم قائمة بالنتائج المحتملة لكل حالة من الحالات ، وبالتالي نتائج هذا التحليل يكون فرضيات لما سيحدث مستقبلاً .

- **تحليلات إرشادية توجيهية Prescriptive Analytics** : الكشف عن الإجراءات التي يجب إتخاذها مستقبلاً وهذا النوع هو الأكثر قيمة لأنه يعطيك القرار وليس المعلومة فقط وهذا هو أقصى طموح وصلت إليه سابقاً تحليلات البيانات ، حيث تستخدم نتائج التحليلات الوصفية والتشخيصية والتنبؤية وتضيف إليها إقتراحات بناء على القرارات التي تم إتخاذها سابقاً في المؤسسة.

وبعد استعراض أهم ما جاء فى تعريف تحليل البيانات ، تبدأ الباحثة فى عرض ومناقشة المحاور الرئيسية للموضوع محل الدراسة وفقا لما جاء فى المصادر المدروسة :

<u>طبيعة البيانات الواجب تحليلها</u>

بعد الإطلاع على أدبيات الموضوع وجد أنه تم تقسيم طبيعة هذه البيانات من وجهات نظر مختلفة ، ففي **أدبيات الموضوع باللغة العربية و الإنجليزية** ، تم معالجة هذا المحور من خلال أربع محاور فرعية وهم : البيانات كمعلومة – كتكنولوجيا – كأسلوب – كتأثير، ويتم استعراض ذلك كالتالي :

عرف (Akoka2017) ، (Keil2014) ، (Suthaharn2014) ، (أحمد خيري ٢٠١٨) بأنها مجموعات البيانات التي لايمكن إلتقاطها ومعالجتها من قبل أجهزة الحواسيب التقليدية ، وأنها أصول المعلومات التي تتسم بالحجم العالي والسرعة والتنوع ، وتتسم بأنها معقدة جداً وبالتالي يجب إستخدامها بكفاءة وفعالية . وأضاف (أحمد فايز سيد ٢٠١٨) بتقدير كمي لهذه البيانات بأن حجمها يصل إلى المئات من التيرابايت أو البيتابايت أو أكثر للحزمة الواحدة من البيانات .

وأضاف (Cuzzocrea2014) بأن طبيعة هذه البيانات تختلف عن البيانات الآخرى ، حيث أن ليست كل هذه البيانات مهيكلة Structured Data ويمكن وضعها في جداول وإدارجها في برامج التحليل

كالأرقام ، ولعل الصور والفيديو خير دليل على ذلك ، وبالتالي هذه البيانات غير المهيكلة Unstructured Data لا يمكن تنسيقها في جداول مثل البريد الإلكتروني ،واتفقت مع ذلك (خلود بنت خالد ، هاجر سليمان ٢٠١٨) لكن أضافت نوع آخر ألا هو بيانات شبه مهيكلة وهي أقرب للمهيكلة ولكنها تفتقر إلى البنية التنظيمية كالجداول وقواعد البيانات ، وأن ما يميز المهيكلة عنها هي إمكانية استخراج المعلومات منها مثل قواعد بيانات Nosql . واختلفت (أسماء بنت سعيد ، هالة بنت خميس ٢٠١٨) في طبيعة هذه البيانات وإقتصرت على البيانات الاحصائية فقط كنوع من أنواع هذه البيانات الواجب تحليلها حيث تحويل البيانات الخام إلى مؤشرات وأرقام إحصائية يتم تمثيلها في شكل صور أو رسوم بيانية أو جداول لتصف ظاهرة أو موضوع معين . كما اقتصر أيضاً (على بن ذيب ٢٠١٧) على أن هذه البيانات متمثلة فقط في البيانات غير المهيكلة المخزنة بطريقة غير منظمة لا تجعل الاستفادة منها ميسوراً .

وعالج هذا المحور (Bizer2011) ، (Madden2012) ، (محي الدين نجيب ٢٠١٨) ، (يوسف على البكري ٢٠١٨) بإضافة خمس سمات لهذه البيانات : -

- **الحجم volume :-** حجم هذه البيانات الواجب تحليلها كبير جدا بالنسبة للحواسيب الشخصية ، فدقيقة واحدة من فيديو غير مضغوط يمكن أن يستهلك ما يقرب من جيجا بايت على مساحة القرص ، و البيانات المتعلقة بالصناعة مثلا تتضخم لدرجة انه من المستحيل تحليلها باستخدام كمبيوتر واحد ، فبرمجيات الحاسب الآلي يمكن أن تعالج بيانات بقدر ميجابايت من وحدة التخزين ، أما البيانات الخاصة بالتيرابايت و البيتا بايت يتطلب أدوات خاصة.
- **السرعة velocity :-** الخاصية الثانية الرئيسية هي السرعة ، في عصرنا الرقمي ، تولد المدونات على سبيل المثال بيانات بمساحات عالية في أسرع وقت ، تويتر مثلا بمثابة مدونة صغيرة ويعتبر مثال جيد على سرعة البيانات الضخمة فإنه ينتج حوالي ٦٠٠٠ تدوينه كل ثانية ، و بالتالي حوالي ٣٥٠,٠٠٠ تدوينه في الدقيقة ، و ٥٠٠ مليون تدوينه في اليوم وحوالي ٢٠٠ بليون تدوينه في السنة . وأضافت (علا نبيل٢٠١٩) بأن السرعة هنا تدل على سرعة توليد البيانات وليس ذلك فحسب ولكن سرعة معالجتها)
- **التنوع variety :-** تتنوع البيانات ما بين بيانات نصية مثل المدونات و نصوص الرسائل ، و بيانات غير نصية مثل الفيديو ، الصور و التسجيلات الصوتية فيمكن للبيانات التقليدية أن تسجل داخل ملف أكسس ، في حين أن البيانات الضخمة تحفظ في أماكن مختلفة (المدونات ، رسائل الدردشة ، زيارات المواقع ، البريد الالكتروني) وهذه البيانات لا يمكن إدخالها على ورقة ملف أكسل .

<u>وأضاف على هذه الخصائص أو السمات (Madden2012) كل من :</u>

- **القيمة value :- قيمة البيانات الضخمة أنها تنقسم إلى عنصرين وهما :**

  أولا :- تكلفة التجهيز و إنشاء البنية التحتية .
  ثانيا :- قيمة إتاحة و تصفح البيانات .
  فتنفق الشركات مبالغ ضخمة لازمة لبناء البنية التحتية اللازمة لتخزين و معالجة تلك البيانات أي (العنصر الأول) ثم تكسب أموالاً بعد التصفح و الإتاحة أي (العنصر الثاني)
- **الدقة veracity :-** يقصد بها ضبط البيانات ، فالبيانات التي لا تتصف بأي من تلك الخصائص السابقة سواء حجم ، سرعة ، تنوع أو قيمة تزيد من تكاليف البنية التحتية و تقلل من القيمة النهائية المكتسبة بعد إتاحتها.

وأضافت (علا نبيل٢٠١٩) كل من :

التصور / الرؤية Vision

فالبيانات لا جدوى لها دون تصورها لفهم طبيعتها بشكل أوضح والمقصود بتصورها أن يتم تحويل البيانات المسردة نظريا الى رسومات تعبر عنها سواء رسومات بيانيه أو إحصائيات .. الخ ، ولا بد من وضع رؤية خاصة بطبيعة البيانات التي يتم تجميعها خصوصا أثناء القيام بمشروعات معتمدة على تلك البيانات حيث أنه يوجد الكثير من البيانات لا جدوى لها .

التحقق Verification

تلك الخاصية معتمدة على الخاصية السابقة ، فان لم يتم العمل بشكل جدي على الخاصية السابقة ، لا تقوم البيانات الضخمة بدورها في التحقق من أمور ما و توضيح الرؤى الخفية لاتخاذ القرارات السليمة .

الفاعلية Validation

تحليل البيانات الضخمة بعد تصفيتها وعزل غير المرغوب فيه أمر لابد منه ، حيث أن تحليل تلك البيانات توضح الكثير من الأسئلة في ذهن العاملين مع تلك البيانات وعليها ، فان لم تحقق تلك البيانات فاعليتها بعد تحليلها في الأمر المرغوب فيه ، فهي بالتالي لا تعد بيانات ذات كفاءة.

وتناول التعريف بطبيعة هذه البيانات المحللة (Gantez and Reinsel) ، (أسماء بنت علي السلامية ٢٠١٨) من وجهة نظر مختلفة عن ما سبق سردة ، حيث عرف هذه البيانات بأنها جيل جديد من التكنولوجيات صمم لاستخراج قيمة إقتصادية من كميات كبيرة جدا من مجموعة واسعة من البيانات وذلك عن طريق التقاطها واكتشافها وتحليلها .

وأضاف (Ward and Baker2013) ، (Manyika2011) بأنها بيانات معقدة جدا ذات سرعة عالية ، وتتطلب تقنيات متقدمة لتمكين التقاطها وتخزينها وإدارتها وتحليلها ، واتفق معهم (محمد أحمد ثابت) في هذا التعريف مع إضافة أنه يمكن اكتشاف الأنماط الخفية بعد تحليلها .

ومن زاوية آخرى استعرض (Chen2012) ، (Bayer2012) ، (Manyika2011) أن طبيعة هذه البيانات المحللة تتطلب مجموعة من الأساليب والأدوات للتعامل معها ، وأن هذه الأساليب والأدوات تتطلب توفير بنية تحتية قوية للقيام بعملية التحليل بشكل سليم وفعال ، وأضاف (بوعناقة سعاد ٢٠١٨) على هذا التعريف المصادر المتعددة للحصول على هذه البيانات كمستشعرات الفضاء ، الهواتف الذكية ، التكنولوجيا القابلة للارتداء .

كما تناول كل من (Rodriguez, Manzahu2016) ، (Chen2012) مدى تأثيرهذه البيانات المحللة حيث قدرتها على نمو الاقتصاد العالمي عن طريق تعزيز الإنتاجية والقدرة التنافسية للمؤسسات ، حيث التأثير في عدة مجالات منها الطاقة والتعليم والصحة و المشاريع العلمية الضخمة أبرزها مشروع الجينوم البشري والذي يحتوي على ٢٥ الف جين الذي تحتوي بدورها على ٣ مليار زوج من القواعد الكيميائية المكونة ل DNA.

وأضاف على هذه المجالات (Archena and Mary2015) بتطوير الخدمات الحكومية عن طريق رصد مدى رضا المواطنين عن الخدمات المقدَّمة لهم وعلى ضوء النتائج المحللة يمكن استنتاج ما يلزم عمله للتطوير والتحسين ، واستعرض أيضا مدى تأثير تحليل هذه البيانات على الصحة حيث كشف وتشخيص الأمراض بناء على تحليل الصور ، حيث تكمن الفكرة في تعليم الحواسيب كيفية التعرف على مختلف أنواع الجروح و الأمراض والتشوهات ، وذلك من خلال عرض مئات من صور الأشعة السينية والرنين المغناطيسي و الأشعة المقطعية وغيرها .

وأضافت (أمل حسين ٢٠١٨) بأنها توفر فرصاً كبيرة لقطاعات التعليم وقطاع الأموال وتكنولوجيا المعلومات والاتصالات والتنمية المستدامة .

كما كون (عمرو حسن فتوح٢٠١٨) علاقه بين تأثير تحليلات البيانات وصنع القرار حيث عرف هذه البيانات بأنها يصعب الاستفادة منها إلا بوجود تقنيات تحليل ذات كفاءة عالية تسمى " تحليلات البيانات الضخمة " تعمل على تحليل وتنظيم وربط هذه البيانات وإتاحتها على هيئة نتائج وحقائق بأقل مجهود للاستفادة منها في صنع القرار .

وإتجه (علي بن ذيب ٢٠١٧) إلى تعريف هذه البيانات عن طريق تصنيفها حسب الصيغ فهناك بيانات نصية ، صوتية ، مقاطع فيديو ، رسوم وأشكال ، وحسب موضوعها الذي تنتمي إليه من موضوعات المعرفة البشرية سواء في مجالات نظرية أو تطبيقية ، وحسب المصدر الذي وردت منه مثل مؤسسة حكومية محلية أو دولية ذات فرع أم عدة فروع ، حسب التاريخ (الإنشاء أو التخزين) .

يمكن تلخيص ما سبق في هذا المحور في شكل جدولي :-

المحاور الفرعية	الدلالة	الدراسات
معلومة Information	وحدات بيانات معقدة	Akoka(2017) , Keil(2014) , Suthaharan(2014)(٢٠١٨) ، أحمد خيري
	زيادة في حجم البيانات المهيكلة وغير المهيكلة وشبه المهيكلة	Larkou(2016) , Andrew,Wong(2015) هاجر سليمان (٢٠١٨) ، علي بن ذيب (٢٠١٧)
	كمية وهيكل البيانات المحللة البيانات الاحصائية	Cuzzocrea(2011) أسماء بنت سعيد ، هالة بنت خميس ٢٠١٨
	إضافة بعض السمات للبيانات المحللة	Bizer (2011) ، محي الدين نجيب (٢٠١٨) يوسف علي البكري (٢٠١٨) ، علا نبيل (٢٠١٨)
	الخصائص والبنية التحتية مصادر البيانات	Madden(2012) ، بوعناقه سعاد (٢٠١٨)
تكنولوجيا Technology	جيل جديد من التكنولوجيا ، مصمم لاستخراج قيمة إقتصادية من البيانات المحللة	Gantez and Reinsel(2011)
	مجموعة من البيانات المعقدة باستخدام سلسلة من التقنيات	Ward and Baker (2013) Manyika (2011)
الأساليب Methods	تطبيق مجموعة من التقنيات المستندة إلى البنية التحتية لتوليد الحقائق .	Chen(2012), Bayer(2012) , Manyika(2011)
التأثير Impact	تأثير ناتج تحليل هذه البيانات في مجالات متنوعة مثل الكيمياء الحيوية وعلم الوراثة	Rodriguez , Mazahu (2016) Chen(2012)
	تأثير تحليلات هذه البيانات في مجال الصحة والحكومة	Archenaa and Mary(2015)
	صنع القرار	حسن فتوح (٢٠١٨)
	التعليم والتنمية المستدامة	أمل حسين (٢٠١٨)

وبناء على عرض الجدول السابق الموضح لمحاور أدبيات الموضوع حول طبيعة هذه البيانات المحللة خلال أربعة محاور فرعية وهما (المعلومة – التكنولوجيا – الأسلوب – التأثير) ، لذلك يمكن وضع تعريف يشرح العلاقة بين تحليلات البيانات والمكتبات بأنها تطبيق أساليب تقنية وتحليلية على البيانات المهيكلة في المكتبة مثل عدد أوعية المعلومات الواجب تصنيفها أو فهرستها والبيانات الآخرى غير المهيكلة مثل بيانات تفاعل المستفيدين مع ناتج العمليات الفنية للمصادر والمتاح في شكل خدمات معلومات خلال الموقع الالكتورني للمكتبة ، وذلك وصولا إلى معلومات قيمة يتم اتخاذ قرار فعال بناء عليها لتحقيق التميز التنافسي لمؤسسات المكتبات والمعلومات .

تقنيات تحليلات البيانات Data Analytics Techniques

يمكن تحليل البيانات المعقدة باستخدام العديد من التقنيات واقترح (Rotriguez 2016) ، (Maniyka 2011) تطبيق تقنيات الرياضيات ، والتي غالباً ما تستخدم في صنع القرار حول نوع العلاقة الموجودة بين المتغيرات وقد تنتج هذه العلاقات ع بعض العلاقات السببية الأساسية .

ومن الأساليب الآخرى المستخدمة على نطاق واسع والتي إقترحها (BU etal 2012) ، (أحمد خيري ٢٠١٨) هي التعلم الآلي ، والغرض من هذه التقنية هو نقل البيانات بطريقة تمكن من المساعدة في التنبؤ بأي بيانات مخفية . وفي الوقت الحالي يتم استخدام التعلم الآلي على نطاق واسع في السوق والمبيعات ويتبنى نظام لاقتراح المستخدمين المستهدفين ، كما يستخدم من قبل العلماء والباحثين في مختلف مجالات العلوم لتفسير هذه البيانات والحصول على المعرفة خاصة في بعض المجالات مثل علم الأحياء والفيزياء . وأضاف (أحمد خيري ٢٠١٩) تقنيات معالجة اللغة الطبيعية (NLP) وهي شكل من أشكال الذكاء الاصطناعي في مجال علوم الحاسوب واللغويات تهتم بالتفاعلات بين الحاسوب واللغات الطبيعية ، فعلى سبيل المثال تتعامل المؤسسات الصحية مع كم كبير من البيانات والمعلومات سواء عن المؤسسة أو عن المرضى أو عن الأطباء مثل براءات الاختراع والأبحاث التنافسية ، ونظراً لان هذه الأنواع من البيانات تتكون إلى حد كبير من اللغة ، وبالتالي معالجة اللغات الطبيعية لهذه البيانات تتيح فرصة ذهبية للاستفادة مما هو متاح في المؤسسات للكشف عن الأنماط والاتجاهات عبر مصادر البيانات المتباينة .

وأيضاً تقنية التنقيب عن البيانات Data Mining كواحدة من تقنيات الحوسبة الذكية التي تعمل في سياق إدارة بيانات الويب كوسيلة تهدف إكتشاف البيانات واستخراجها بغية تحويل الويب من وضعه الراهن إلى شبكة معرفية (مؤمن النشرتي ، ٢٠١٤) وأضاف كل من (Sgueza , Squicaela , Avila ordouez , Cottrysse 2015) أنها تساعد على تحديد الأنماط السلوكية بين مستخدمي المكتبة وأنماط الموظفين وأنماط استخدام موارد البيانات قي المكتبة باستخدام Bibliomining وينتهي بملخص شامل عن كل ذلك لتحديد مجالات القصور والتنبؤ باحتياجات المستقبل .

وأضاف (AL daihania and Abrahams 2016) تقنية تحليل النصوص Text-mining لتحليل النصوص المتعلقة بوسائل الاعلام الاجتماعية للمستفيدين وتضمين ذلك من العمليات الفنية الأساسية للمكتبة ، والذي يتم بناء عليه تقديم خدمات فعالة .

وتقنيات معالجة الإشارة A signal Process تقنية آخرى لتحليل البيانات المعقدة ، حيث تبسط نطاق المعرفة المتزايدة في العديد من مجالات العلوم من خلال تطبيق عدة أنظمة رياضية (Slavakis , Matoes 2014) .

ومن التقنيات أيضاً تقنية المرئيات Visualization حيث تعمل على تحويل البيانات إلى شكل مرئي ، بحيث يمكن تحليل خصائص هذه البيانات أو العلاقات بينهم ، وتعد واحدة من التقنيات الأكثر قوة وجاذبية لاستكشاف البيانات . (Han etal 2011 ، (أحمد خيري ،٢٠١٨) ،(علا نبيل ٢٠١٩) ، في حين

استفاضت (علا نبيل ،٢٠١٩) في استعراض هذه التقنية حيث التعريف وأنواع التمثيل المرئي المناسب لكل نوع من أنواع البيانات في المكتبات سواء بيانات مصادر معلومات أو بيانات خاصة بالأشخاص (مستفيدين أو امناء المكتبات) أو بيانات متعلقة بدورات وورش تدريبية .. الخ ، ونهاية رصد أهم المكتبات المفتوح المصدر للمرئيات مثل Tableau ، Zing Chart ، Timeline .

كما صنف (Alejandro Santovena, 2013) تقنيات تحليلات البيانات إلى :-

- K- Nearst Neighboars : تعد تقنية للتصنيف والتنبؤ حيث تسمح بتقسيم البيانات إلى مجموعات صغيرة ، وكل مجموعة تتألف من عدد من السجلات مماثلة لبعضها البعض .
- Naïve Bayes : جاء لكل مشاكل تصنيف البيانات عن طريق الاحتمالية حيث تعمل على احتمالية الحدث قبل وقوعه .
- شجرة التصنيف والانحدار Classification and Regression Tree : تستخدم لفصل السجلات Records في شكل مجموعات فرعية شجرية ، ويعتبر هذا الأسلوب الأنسب حين وجود مجموعة واسعة حيث يوفر قاعدة تصنيف سهلة الفهم .
- التحليل التمييزي Discriminant Analysis : تميل إلى استخدام أكثر من طريقة للتصنيف الاحصائي للبيانات ، وذلك لتمييز الجوانب المختلفة للفئات في مجموعات البيانات Dataset
- الانحدار اللوجستي Logistic Regression : يستخدم لتصنيف سجل جديد لفئة غير معروفة في واحدة من الفئات الموجودة بالفعل استناداً إلى متغير التوقع التنبؤي .
- الشبكات العصبية Neural Networks : تجمع هذه التقنية بين المعلومات بطريقة مرنه للغاية والاعتماد على هذه العلاقات المعقدة بين المتغيرات مما يؤدي إلى التنبؤ بالحدث قبل وقوعه ، وتتضمن هذه الشبكات ثلاث طبقات [طبقة الادخال – الطبقة المخفية – طبقة المخرجات] وتدفق المعلومات يأتي في إتجاه واحد.
- التحليل الجغرافي Spatial Analysis : مزيج من التقنيات الاحصائية والجغرافية والهندسية ، تأتي غالباً لتحليل البيانات المكاني من نظم المعلومات الجغرافية (GIS) حيث إدراج وإلتقاط بيانات المواقع والعناوين لفهم رغبة المستفيد في شراء منتج مرتبط بمكان معين .

وأجمع كلاً من (Cheistine M.shupala 2012) ، (Elham Abdel-razik 2012)على استخدام تقنية تغليف البيانات Data envolepment analysis في تحليل بيانات عينة من الجامعات بالاعتماد على بيانات الميزانية ، أعضاء هيئة التدريس ، الخريجين ، الطلاب الحاليين و استمثار نتائج ذلك في صنع القرار وإتفقوا أيضا بأن هذه التقنية تستخدم على نطاق واسع في تقييم آداء وحدات صنع القرار في وحدات الأعمال والوكالات الحومية وإدارات الشرطة والمستشفيات والمؤسسات التعليمية .

كما إتفق كلاً من (علي بن ذيب الأكلبي ٢٠١٧) ، (أحمد فايز أحمد ٢٠١٩) ،(علا نبيل ٢٠١٩) ، (أحمد خيري ٢٠١٨) على ظهور عدد من النظم التي إختصت بأعمال تحليل هذه البيانات وأنها بمثابة تقنيات وأدوات برمجية مثل Hadoop ونظام Sap Hana حيث تساعد هذه التقنيات مستخدميها في المؤسسات التجارية وغير التجارية التي تقتنيه من إنجاز العمليات التجارية بشكل أسرع من خلال ما توفر لها من بيانات معالجة تساعد في إتخاذ القرار . واضاف كل من (أحمد فايز ٢٠١٩) ، (علا نبيل٢٠١٩) تقنيات معالجة البيانات التشغيلية الكبرى وهي النظم التي توفر قدرات تشغيليه للأعمال الزائدة التفاعلية في وقت مناسب لها ، حيث يتم التقاط البيانات وتخزينها بشكل أولي ، ومثال على هذه النظم Monog DB

حيث الاستفادة من الحوسبة السحابية لاتاحة إمكانية تشغيل الحاسبات الضخمة بشكل رخيص فعال ، مما يسهل إدارة كم هائل من البيانات وتحليلها مع حد أدنى من التكويد وبدون الحاجة إلى علماء البيانات ، وأيضا تقنيات نظم قواعد بيانات المعالجة المتوازية Map Reduce حيث يوفر تحليل البيانات المكملة للامكانات التي توفرها SQL ويمكنه توسيع نطاقه من خادمات فردية إلى الآلاف من الخادمات الافتراضية . وتخصص كلاً من (أحمد خيري ٢٠١٨) ، (علا نبيل ٢٠١٩) في التعريف بقواعد بيانات NOSQL في تحليلات وإدارة بيانات المكتبات وهي عبارة عن قواعد بيانات تتيح تخزين واسترجاع البيانات بعد نمذجتها بتقنيات تختلف عن تلك المستخدمة في نمذجة البيانات في قواعد البيانات العلائقية ، ويميل المطورون Developers إلى استخدام هذا النوع من قواعد البيانات لما تقدمة من بساطة في التصميم وقابلية للتوسع ، لكن أضافت (علا نبيل ٢٠١٩) منهجية تخزين البيانات في NOSQL وأشهر أنواع قواعد البيانات وأهم الشركات المستخدمه لها .

إذن تزداد أهمية تحليلات البيانات بسرعة ، وبالتالي لابد من الاستعداد من قبل المؤسسات للاستفادة من أصول البيانات الخاصه بهم لتحقيق ميزة تنافسية ، وذلك من خلال ما تقدمه هذه التحليلات من المرونة في الأداء ، ويلخص الجدول التالي موضوعات تقنيات تحليلات البيانات في الأدبيات كالتالي :-

المحاور الفرعية	الدلالة	الدراسة
التقنيات الاحصائية Statistics Techniques	تستخدم لاتخاذ القرارات حول نوع العلاقة الموجودة بين المتغيرات	Manyika(2012)
تقنيات التعلم الآلي Machine Learning Techniques	استخدام الاحصائيات بشكل معين لتوقع أو شرح المعلومات الخفية	Bu etal(2012) Hamdy mousa (2013) أحمد خيري (٢٠١٨)
تقنيات معالجة اللغة الطبيعية NLP	الكشف عن الأنماط والاتجاهات عبر مصادر البيانات المتباينة	أحمد خيري (٢٠١٨)
تقنيات التنقيب عن البيانات Data Mining	إكتشاف البيانات واستخراجها بغية تحويل الويب إلى شبكة معرفية .	Han etal(201) مؤمن النشرتي (٢٠١٤)
	استخدام Bibliominig لتتبع الأنماط السلوكية والتغييرات	Siguenza -Guzman (2015)
	إتباع تقنية تحليل النصوص Text-mining لتحليل وسائل الاعلام الاجتماعية للمكتبات	AL Daihani and Abrahams(2016)
تقنيات تغليف البيانات DEA	تحليل بيانات وحدات الأعمال	Cheistine m.shupala (2012) , Elham Abd el-razik (2012)
تقنيات معالجة الإشارة Signal Processing Techniques	خوارزميات إحصائية مستندة إلى معلومات تهدف إلى تبسيط تحليلات البيانات في الوقت الحقيقي	Slavakis (2014)
تقنيات المرئيات Visualization	عرض توزيع عالي من الاحصائيات من خلال العلاقات والخصائص المشتركة لمجموعة من البيانات	Ma etal (2012) أحمد خيري (٢٠١٨) ،علا نبيل (٢٠١٩)

يشير الجدول السابق إلى خمس مواضيع مرتبطة بتقنيات تحليلات البيانات الموجودة في أدبيات الموضوع { تقنيات إحصائية – تقنيات التعلم الآلي – تقنيات معالجة اللغة الطبيعية – تقنيات التنقيب عن البيانات – تقنيات معالجة الإشارة – تقنيات المرئيات } ، وتوصي الدراسة بتطبيق هذه التقنيات على مؤسسات المكتبات ومراكز المعلومات ، حيث الاعتماد على التقنيات الاحصائية وتقنيات التعلم الآلي والتنقيب عن البيانات وتقنيات معالجة اللغة الطبيعية لاستخراج احصائيات حول مجموعات المكتبة مثل عدد المصادر وما تم استبعادة وتنقيته على سبيل المثال ، وعدد المصادر المتاحة على الموقع نص كامل ومدى إقبال المستفيدين على هذه المصادر .. الخ

وفي المرحلة التالية يتم الاعتماد على تقنيات المرئيات بحيث يتم التعبير عن كل هذه الاحصاءات بشكل مرئي يناسب طبيعة كل بيان سواء مهيكل أو غير مهيكل أو شبه مهيكل وصولاً إلى تحديد نقاط القوه والضعف واتخاذ القرار المناسب .

أدوات تحليلات البيانات Data Analysis Tools

تنويه : تم ذكر Hadoop – Sap Hana – Monog DB - NOSQL كتقنيات لتحليلات البيانات في الإنتاج الفكري العربي ، ولكن بعد الاطلاع على الانتاج الفكري الأجنبي وجد أن هناك فرق بين الأدوات والتقنيات وتم الفصل بينهم في استعراض الأدبيات ، وبناء على ذلك سوف يتم إضافتهم في الجدول الخاص بالأدوات والتي سيتم إدراجه بعد مناقشة الأدوات كاملة كالتالي :-

صنف (Chen 2014) أدوات تحليلات البيانات إلى ثلاثة أنوع من النماذج تماشياً مع نوع التحليل كالتالي :-

١- النوع الأول (Batch Analysis) حيث يتم تصفية البيانات أولاً ثم تحليلها .
٢- النوع الثاني (Steam Analysis) حيث تحليل البيانات في الوقت الحاضر لاستخلاص نتائجها
٣- النوع الثالث (Interactive Analysis) حيث يتعلق بمعالجة البيانات باستخدام التحليل التفاعلي مما يسمح للمستخدمين بالقيام بأنفسهم بعملية تحليل البيانات .

النوع الأول Batch Analysis

يشمل هذا النوع – Google Map – Map Reduce – Apache Hadoop – Microsoft Dryad Apache Mahout

Google Map Reduce : هو إطار عمل يمكن من كتابة تطبيقات لمعالجة كميات ضخمة من البيانات باستخدام المعالجة المتوازية في تجمع حاسبات كبيرة ، فهو نموذج للحوسبة الموزعة تعتمد على تقنيتي Map و Reduce بحيث يتم العمل في ثلاث مراحل :-

١- الخريطة Map : تقسيم البيانات إلى أجزاء
٢- ترسل مخرجات Map إلى مرحلة الترتيب
٣- تقوم تقنية التقليص Reduce بأخذ مخرجات المرحلة السابقة كمدخلات لتجميعها .

(Zhifeng 2013) ، وعلى الرغم من نجاحها إلا أنه يظهر العديد من القيود مثل الخوارزميات المكررة والنفقات الزائدة (Lopez,delrio,benitez&herrava 2014)

Apache Hadoop : برنامج مفتوح المصدر بإطار تخزين موزع لمجموعات كبيرة من البيانات فيما يسمى بالعناقيد الحاسوبية Computer Clusters ، مما يعطية خاصية التوسعية Scalability و الاعتمادية Reliability والعمل دون قلق من فشل الأجهزة ، حيث يتزود بشبكة Network عالية السرعة

إلى جانب الاحتفاظ بالبيانات بأكثر من نسخة في أكثر من مكان ولدية القدرة على تخزين ومعالجة بيانات ضخمة من جميع انواع البيانات مع القدرة على التعامل مع وظائف متزامنة (Begoli,hery 2012) ، (Boker 2012) . وإتفق (Mohamed malhat etal 2019) في ذلك ووصفها بأنها أداة لتقسيم البيانات ، ففي حالة أن الأساليب التقليدية أصبحت غير قادرة على معالجة البيانات المعقدة فجاءت هذه الأداة لحل هذه المشكلة حيث تقسيم مجموعات البيانات إلى مجموعات فرعية يمكن إدارتها وتطبيق الاساليب التقليدية على كل مجموعة فرعية على حدة ، ولكن التحدي هنا يظهر في حالة زيادة عدد المجموعات الفرعية للبيانات المقسمة .

النوع الثاني : Stream Analysis

تستخدم على نطاق واسع في التحليل المالي ، والتداول عبر الانترنت والاختبارات الطبية ، وذلك يلزم وجود أطر نظرية وتقنية فعالة لدعم هذه التقنية (Wu,zhu 2014) ، وهذا يشمل معالجة كميات كبيرة من البيانات واستخدام عدة أجهزة لموازنة عبء العمل ، وللحصول على نتائج فعالة لابد من تطبيق أربع أدوات وهما :-

Apache Storm : مصدر مفتوح لديه العديد من التطبيقات التي تعمل في وقت متزامن ويستخدم مع أي لغة برمجة ، وعلاوة على ذلك تعد أداة سريعة حيث تتمكن من معالجة أكثر من مليون عقدة من البيانات في الثانية (Rodriguez – Mazahua 2016) .

Apache S4 : منصة تسمح للمبرمجين من تطوير التطبيقات بسهولة للمعالجة المستمرة لتيارات غير محدودة من البيانات ، وفي هذا النظام تقوم عناصر المعالجة بإرسال الرموز التي تمكن من الاتصال فيما بينها في شكل أنشطة البيانات Data Activities ، وهذه العناصر تمتاز بالشفافية والكفاءة ، ويمكن لمنشئ النظام كتابة هذه العناصر في لغة برمجة java (Neumeyer , Robbins & Kasars 2010)

Apache spark : يعد محركاً سريعاً لمعالجة تحليلات البيانات على نطاق واسع ، مع واجهات تسمح لمعالجي البيانات بتنفيذ العمليات ، ويعد الأكثر نشاطاً في مجتمع المصادر المفتوحة ويستخدم العديد من اللغات مثل Java , Paython (Stocia 2014) .

MOA : برنامج مفتوح المصدر يستخدم في تفسير البيانات وتصنيفها حيث يتيح كل من التنقيب عن البيانات Data mining و محاكاة عالم الرسم البياني Graphic mining ، ويتكامل هذا البرنامج مع برنامج آخر يسمى ST و Typhoon لتنفيذ مهمة التحليل (Fan & Bifet 2012) .

وعلى نحو أكثر تحديداً ، طبق كلاً من (Al daihani and Abrahamas 2016) نهجاً لاستخراج النصوص ، وذلك للتحقيق في مجموعة كبيرة من التغريدات TWEETS الخاصة بالمكتبات الأكاديمية ، وأبرزت النتائج المتعلقة باستخدام طريقة تحليل المحتوى في تقييم معلومات التواصل الاجتماعي للمكتبات ، وظهور مدى أهمية هذه البيانات المحللة في شكل احصاءات ساعدت المكتبة في التخطيط الاستراتيجي في المستقبل واتخاذ القرارت المثمرة .

النوع الثالث : Interactive Analysis

نظراً للأدبيات السابقة ، تساعد الأدوات التحليلية التفاعلية الثلاثة التالية أنظمة البيانات المعقدة :-

Apache Drill : إطار يدعم تحليل تفاعلي لكميات كبيرة جداً من البيانات تزويدها بإحصائيات عديدة (Rodriguez – Mazahua etal 2016) .

SpagoBI : أداة تستطيع التعامل مع كميات ضخمة من " البيانات غير المهيكلة " وتستخدم على نطاق واسع مع البيانات التجارية ، وتعطي معاني قيمة بعد تفسيرها حيث توفر حقائق ذات معنى تؤثر في إتخاذ القرارات ، وذلك باستخدام أدوات مثل Self – service BI ولوحات عمل وتقارير ..الخ ، كما أنها من الممكن أن تدمج مع محرك OBE لمساعدة المستخدمين إنشاء تحليلاتهم مباشرة مع مجموعات البيانات الخاصة بهم (Franceschinis 2013)

لوحظ خلال عرض الأدوات الخاصة بتحليلات البيانات في أدبيات الموضوع باللغة الانجليزية ، بعدم حصرهم بشكل كاف ، حيث يوجد الكثير من الأدوات الآخرى مثل :-

Plotly : تقدم تلك المكتبة المرئية أدوات لتصور البيانات في شكل رسوم بيانية Charts وعروض Presentation ولوحات عمل تفاعلية Dashboard ، ويمكن إجراء تحليلات البيانات بإستخدام JavaScript أو Python أو Excel وتقدم العديد من الخيارات لإستيراد البيانات Importing Data .

Tableau : تقدم العديد من الأشكال المرئية كالرسوم البيانية Graphs والمخططات Charts والخرائط Map وغيرها ، وتلك الأداة متاحة بشكل مجاني وتمكن المستخدم من سحب بياناته و إدراجها داخل النظام والتعامل معها ومشاهدة التحديثات الخاصة بتلك البيانات في الوقت الحقيقي Real-Time

Timeline: تقدم تلك الأداة مرئيات مطورة بحيث تغطي الكثير من المعلومات في مساحة معينة مضغوطة في شكل عناصر ، وكل عنصر Element يظهر المعلومات المتعلقة حولة بمجرد الضغط عليه . (علا نبيل ٢٠١٩) .

الموضوع	الأداة	الدلالة	الدراسات
Batch Analysis التحليل بالدفعات	Google map reduce	معالجة البيانات في عناقيد ضخمة	Zhifeng and yang (2013) Hashen etal (2015)
	Apache Hadoop	بنية تحتية ومنصة	Begoli and Herey (2012) Boker , carey (2012)) Mohamed malhat etal 2019
	NOSQL	قواعد بيانات غير علائقية	أحمد خيري (٢٠١٨) ، علا نبيل (٢٠١٩)
	Sap Hana	تحليل البيانات التجارية	أحمد فايز (٢٠١٨) ،علي بن ذيب(٢٠١٧)
	Monog DB	تحليل البيانات مع حد أدنى من التكويد	أحمد فايز (٢٠١٨) ، علا نبيل (٢٠١٩)
Stream Analysis التحليل المتدفق	Apache Storm	آلة الحساب في الوقت الحقيقي	Rodriguez-mazahua etal (2016)
	Apache S4	منصة توزيع الحاسبات	Neumeyar (2010)
	Apache Spark	معالجة البيانات واسعة النطاق	Stocia (2014)
	MOA	إطار التنقيب عن البيانات	Fan and Bifet (2012)

الموضوع	الأداة	الدلالة	الدراسات
Interactive Analysis التحليل التفاعلي	Apache Drill	أداة استعلام لـ Hadoop ، Nosql	Rodriguez-mazahua etal (2016)
	Spago BI	أنظمة تفسير البيانات التجارية	Bostock etal (2011)
	D3.JS	تفاعلي	Franceschini (2013)
	Plotly ,Tableau, Timeline, Zinchart	مكتبات مفتوحة المصدر لعرض البيانات المحللة	علا نبيل (٢٠١٩)

إتجاهات تحليلات البيانات في المكتبات ومراكز المعلومات

لابد من التعرف أولاً على إدارة البيانات وتحليلها قبل إستعراض الإنتاج الفكري في التخصص ، فوفقاً لتقرير(وزارة تكنولوجيا المعلومات والإتصالات بقطر ،٢٠١٤) أنه سيتحول التركيز من كيفية تخزين ومعالجة كميات كبيرة من البيانات إلى كيفية إتخاذ قرارات مدروسة ودقيقة وفعّالة وزهيدة التكلفة ، وقد يتطلب ذلك إحداث نقلة نوعية في كيفية إدارة البيانات و تحليلها ، كما يبين الشكل التالي وزارة تكنولوجيا المعلومات والاتصالات بقطر (٢٠١٤)

شكل رقم (٢) إدارة البيانات قبل وبعد البيانات الضخمة

- **إيضاحاً لهذا الشكل** فبالنظر الي المرحلة الثالثة في ادارة البيانات كانت قبل البيانات الضخمة يتم حفظها في مستودع دون فصلها أو اختيار البيانات ذات الأهمية فقط والتغاضي عن الباقي ، لكن بعد البيانات الضخمة أصبح في تلك المرحلة عدم قدرة المستودعات أو قواعد البيانات على حفظ هذا الكم من البيانات فأصبح ولا بد من تحليلها أولا باستخدام أدوات وتقنيات ملائمة لحجمها .
- وأيضا بالنظر إلى المرحلة الثالثة في التحليل قبل البيانات الضخمة كانت مجرد بيانات فردية لكن بعد البيانات ضخمه أصبح هناك ربط وتواصل بين أكثر من مجموعة من البيانات وذلك لإيجاد علاقة و الحصول على معلومة قيمة لاتخاذ قرار ما مناسب .

وبالتالي يمكن استخراج معلومات قيمة من البيانات الخام ، ولكن ذلك يتطلب شروط ضرورية :-

الدقة accuracy والتأمين insurance وقابلية التوسع Scalability ، حيث أن دقة تقنيات استخراج البيانات تؤثر مباشرة على جودة الاحصائيات الناتجة فمن المهم جداً تحقيق الدقة بقدر الإمكان ، وقابلية التوسع هي صفة مميزة للبيانات المهيكلة وغير المهيكلة وبالأخص غير المهيكلة نظراً لضخامة كمية الحقائق الواجب معالجتها (Khan etal 2017)

ويوجد أربع أنواع من الاحصائيات تحتاج المكتبات لاستخلاصها من البيانات المهيكلة المتولدة داخلها وهي :-

- البيانات الوصفية Metadata
- بيانات المؤلف Writer data
- الاستشهادات Citation (Khan etal 2017) .

ومع زيادة إنتشار وتوليد البيانات وإزدهار تقنيات الحوسبة السحابية ، تحول التركيز من كيفية تخزين ومعالجة كميات كبيرة من البيانات فقط إلى اتخاذ قرارات فعالة وزهيدة التكلفة ، وذلك أدى إلى إحداث نقلة نوعية في كيفية إدارة البيانات وتحليلها . وبالنظر إلى حقل المكتبات والمعلومات ، نجد أنه يتضمن جمع وحفظ ووصف المعلومات ،الاتنظيم الإدارة ، الإسترجاع ، وهذا يوسع الميدان من حقل أو مجال يركز بشكل أساسي على المنشورات العلمية إلى حقل يفحص البيانات التي أنشأها الباحثون والمستفيدون بشكل عام . (Reinhalter & wittmann 2014)

وعلى الأساس الاستكشافي لواقع المكتبات ، قامت كلاً من (خلود بنت خالد ، هاجر سليمان ٢٠١٨) بدراسة واقع استخدام تحليلات البيانات بهيئة الوثائق والمحفوظات الوطنية عن طريق إجراء مقابلات مع قسم النظم والبرمجيات والشبكات في الهيئة ، وتوصلوا إلى أن الهيئة كرست جميع الجهود للتماشي مع كل مايستجد في مجال تكنولوجيا المعلومات وتمثل ذلك في إقتناء هذه البرمجيات والأدوات التحليلية ، وإتفق في ذلك (علي بن ذيب ، ٢٠١٨) عن طريق دراسة واقع نظام إتقان- " وهو عبارة عن نظام للتخطيط وإدارة المعلومات في جامعة الملك سعود لإدارة التميز في الآداء ويسعى إلى توفير مصدر موجه لبيانات الجامعة يساعد على توفير المعلومات المناسبة في الوقت المناسب بجودة عالية " - في جامعة الملك سعود ولكن ركز بشكل أكثر على تأثير هذه التحليلات للبيانات على إتخاذ القرار ووصى بمجموعة من نماذج التحليلات التنبؤية الممكن الاستفادة منها في اتخاذ القرار بجامعة الملك سعود .

واتفق (علي بن ذيب ٢٠١٧) في عملية اتخاذ القرار بالإضافة إلى خلق قيمة مضافة من خلال التحليلات ٠ وأضاف (بوعناقة سعاد ٢٠١٨)، (محي الدين كوكو ٢٠١٨) نقاط قوة وضعف تقنيات تحليلات البيانات بالمكتبات حيث تمثلت نقاط القوة في قدرة تحليلات البيانات على إدارة التغييرات في نماذج البيانات والمعلومات وتصنيفها وتكشيفها وفقا للتقنيات الحديثة ، وتمثلت نقاط الضعف في فقدان المعلومات لقيمتها إن لم تستخدم في الوقت المناسب ، وتعدد أشكال البيانات في المكتبات من منظم وغير منظم ، وزيادة التكلفة المتمثلة في توظيف المحللين وخبراء البيانات وتناقض بعض التحليلات مع بعض أنواع البيانات . واتفق (Alfredo, karen2011) في نقاط الضعف السابقة بالإضافة إلى إيضاح مصادر البيانات الغير المتطابقة المتمثلة في مستودعات البيانات العلمية وأجهزة الاستشعار والشبكات الاجتماعية .. الخ وعدم تجانسها يؤثر على نوع التحليل ، كذلك الحاجة المستمرة لتصفية البيانات Filtering لان ذلك يؤثر على جودة التحليلات النهائية و ذكر أيضاً أن تطبيق تحليلات ذات مغزى يتطلب من المكتبة تحويل هذه البيانات غير المهيكلة في شكل منظم وفقاً لمنهجية محددة تحقق الأهداف المرجوه . واتجهت (أمل حسين عبد القادر ٢٠١٨) إلى ربط تحليلات البيانات بالمكتبة بالمعرفة والاسهام في التنمية المستدامة حيث مواكبة التنمية البحثية والعلمية من قبل المكتبات والاتحادات وجمعيات المكتبات ومؤسسات المعلومات لتنفيذ خطط التنمية المستدامة والسعي نحو تعزيز جهود جمع ورصد البيانات خلال مجتمع المعرفة وتشجيع تقاسم هذه المعرفة التعامل مع الحواجز الرئيسية التي تحول دون الوصول إلى البيانات واستخدامها ووضع استراتيجيات جديدة للتعامل مع تحليلات هذا الكم من البيانات وتضمينه ضمن العمليات الفنية الأساسية بالمكتبات .

واتجهت (آمال بنت سعود العتيبي ٢٠١٨) تطبيق تحليلات البيانات لدور ذات أهمية كبيرة ألا وهو صناعة المعلومات وذلك عن طريق تصنيف وتجهيز البيانات وفقاً للتقنيات الملائمة وإنشاء نظم وقواعد بيانات مختلفة غير تقليدية تساعد على حفظ التسجيلات والبيانات المتعلقة بتكشيف المحتوى للبيانات المهيكلة وغير المهيكلة في نظام واحد وربطهم ببعضهم البعض وإظهار ناتج هذه العلاقة في شكل خدمة للمستفيد . وعلى المستوى الاستكشافي التحليلي تناولت (علا نبيل ٢٠١٩) مشروعات البيانات الضخمة في المكتبات ، بتحليل ثلاث مشروعات وهما Library - Open University Library Data Project AG Data commons Beta – Analytics Toolkit وكانت بمثابة دليل للمكتبات العربية لتطبيق مثل هذه المشروعات وخلال استعراض هذه النماذج تم ذكر مرحلة التحليل والتقنيات الواجب الاعتماد عليها ونوعية البيانات الواجب تحليلها ، في حين إتجه (أحمد خيري ٢٠١٨) إلى دراسة دور مراكز المعلومات الصحية في إدارة البيانات الضخمة وتحليلاتها عن طريق تحليل Swot ووضع تصور مقترح لخطة استراتيجية لتطبيق هذه التحليلات في قطاع الرعاية الصحية في مصر .

وعلى المستوى التقييمي ، سعى (خالد عتيق ، عبدالله بن سالم ٢٠١٨) بقياس وتقييم واقع إمكانات تحليلات البيانات واستخدامها بمكتبات جامعة السلطان قابوس وتوصلت الدراسة إلى أن توافر إمكانات تحليلات البيانات واستخدامها بدرجة متوسطة ووجود صعوبات تواجة الموظفين في التعامل معها .

على المستوى التطبيقي ، قامت مكتبة جامعة Purdue - عند قيام الجامعة بتحليل مستودعها الرئيسي – بالتعاون مع قسم تكنولوجيا المعلومات لطبيق تقنية تلبي احتياجات الباحثين للبيانات المتعلقة بالمجموعات وغيرها من بيانات مهيكلة والآخرى غير المهيكلة مثل تحليل محتوى المنشورات على وسائل التواصل الاجتماعي ، وبناء على ذلك تم الوصول إلى بعض النتائج والاحصائيات الهامة التي عملت على تطوير المكتبة . (Witt 2012)

واتفق معه (Reinhalter & Wittmann 2014) في أن مستودع البيانات في المكتبات هي الطريقة المثالية للسيطرة على الحقائق المنتجة ، وذلك بالطبع بالتزاوج أو التداخل مع الخدمات الاعلامية ، وبالتالي يجب أن ينظر إلى هذه البيانات بأهمية ويتم تضمينها في المكتبات كمصدر هام من مصادر المعلومات يتم جمعه وتصنيفه وتكشيفه واتاحته للمستفيدين .

كما طبق (Young seok , Louise cooke 2017) تحليلات البيانات بالمكتبات العامة في كلاً من مدينتي لندن وسيول باستخدام أسلوب Chernoff Face للتحليل وهو عبارة عن أداة مثيرة للإعجاب حيث يعرض عدة متغيرات متمثلة في وجه الإنسان وبالتالي كل وجه يمثل بيانات شخص واحد أو مؤسسة واحدة أو شركة واحدة أو بلد واحد .. الخ، واعتمد على بيانات الموظفين ، بيانات مجموعات المكتبة ، أعداد المستفيدين والميزانية وعدد مرات الولوج للمكتبة وحجم مبنى المكتبة ، وبالاعتماد على هذه البيانات ولغة البرمجة R لتقييم عمليات المكتبة في في المدينتين حيث الاعتماد على ثمانية متغيرات تم اختيارها استناداً إلى أهميتها وتم تمثيلها كالتالي عرض الرأس يمثل الزيارات وأعلى الرأس يمثل القضايا وحجم الأذن يمثل عدد المكتبات وحجم العين يمثل المجموعات وحجم الأنف يمثل الميزانيات وحجم الفم يمثل عدد الموظفين وإنحناء الفم يمثل عدد الموظفين المهنيين وحجم الوجه يمثل مساحة قاعة المكتبة . واتفق (Lauren Mandel 2016) في تطبيق تقنية التحليل التفاعلي Visualization كتقنية من تقنيات تحليلات البيانات في المكتبات ولكن اعتمد هنا على نظم المعلومات الجغرافية Geograhpic information system GIS ، وذلك لعرض وتصور مرافق المكتبة الجامعية بالاعتماد على تحليل توقيب وأماكن الاستخدام المرافق من قبل الطلاب في فترة زمنية معينة ، وبالتالي تساعد لفهم أفضل حول كيفية استخدام مساحات المكتبة باستخدام هذه الخرائط البصرية الجذابة ، حيث ذكر عمليات التصنيف والتحليل أصبحت غير مقتصرة على أوعية المعلومات فقط بالمكتبة مع هذا التطور التكنولوجي .

وأضاف (Rani 2016) أن كل من المكتبات وأمناء المكتبات مثالية فريدة للعمل مع تحليلات البيانات ، حي٠ث أنها لديها ثقافة مطولة حول معالجة البيانات ، وأن المكتبات مثلها مثل باقي المنظمات ليست بمنأى عن التغييرات التكنولوجية الجديدة وبالتالي لابد من محاكاة تلك الاتجاهات الجديدة التقنية والتي تمكن المكتبات من معالجة نقاط الضعف وتأكيد نقاط القوة . كما أكد ضرورة فهم التغييرات المختلفة لتتمكن من تحقيق الاستفادة القصوى من الموظفين الحاليين وتحديد المتطلبات الجديدة لهم للحصول على المهارات والمعارف الضرورية .

كما أكد (Al- Daihani & Abrahams 2016) أنه يمكن للمكتبات الاستفادة من استثمار تقنيات تحليل النصوص والبيانات في تحليل منشورات مواقع التواصل Posting وبالفعل سوف تحتوي على ثروة هائلة من وسائل الاعلام الاجتماعي في وقت قياسي ، والتي يمكن استخدامها في تعزيز وتطوير المكتبات . وأكد (Heidron 2012) ، (عمرو فتوح ٢٠١٨)، (علا نبيل ٢٠١٩) ضرورة إنشاء قسم خاص بتحليلات البيانات ودعم اتخاذ القرار بالمكتبات لمتابعة جميع ما سبق وتنفيذه .

ويوجز الجدول التالي الموضوعات المتصلة بمجالات الافادة من تحليلات البيانات في المكتبات ومراكز المعلومات حسب أدبيات التخصص

الموضوعات الفرعية	الدلالة	الدراسات
خطط تحليل البيانات المعقدة	إنشاء أقسام تحليلات البيانات في المكتبات	Heidron (2011) عمرو فتوح (٢٠١٨) ،علا نبيل (٢٠١٩)
	بحث مجموعات البيانات الداخلية ، ومعرفة مصادر السجلات الخارجية وربط ذلك لإنشاء علاقات .	Witt (2012)
	إنشاء مستودع بيانات لتحويل البيانات غير المنظمة إلى نموذج منظم ثم تحليلها .	Reinhalter and Wittman (2014)
	تحليل مشروع قائمة بالفعل لتحليلات البيانات في المكتبات	علا نبيل (٢٠١٩)
	خطة لتحليل البيانات بمؤسسات القطاع الصحي في مصر .	أحمد خيري (٢٠١٨)
دراسة واقع تحليلات البيانات في المكتبات	هيئة الوثائق والمحفوظات الوطنية	خلود بنت خالد ، هاجر سليمان (٢٠١٨)
	نظام " إتقان "	علي بن ذيب (٢٠١٨)
	دراسة نقاط القوة والضعف	بو عناقة سعاد (٢٠١٨) ، محي الدين كوكو (٢٠١٨) Alfredokaren (2011)
الفهم النظري والأكاديمي لتحليلات البيانات المعقدة في	العمل مع مجموعات البيانات المعقدة الضخمة وتصنيفها وتكشيف وتحليل محتواها، ومجموعات البيانات ذات الروابط التشغيلية وتحديد الأنماط في	Rani (2016)

الموضوعات الفرعية	الدلالة	الدراسات
المكتبات	الوقت الحقيقي وتوقع المخاطر .	
	فهم التغييرات المختلفة والمهارات المطلوبة	Rani (2016)
دعم الباحثين	اتاحة نتائج تحليلات البيانات مهيكلة وغير مهيكلة بما في ذلك الاحصائيات وغيرها للباحثين في مكتبة كلية الطب الحيوي Biomedical .	Keil(2014)
	إتاحة البيانات المهيكلة ومجموعات البيانات الآخرى غير المهيكلة بتنسيقات مختلفة من صور وفيديو وانفوجرافيك وخرائط ومجموعات بيانات Datasets للباحثين بالمكتبة القومية الزراعية بأمريكا .	علا نبيل (٢٠١٩)
استثمار فرص ممارسة تقنيات تحليلات البيانات	تحليل منشورات مواقع التواصل الاجتماعي للمكتبة ومقارنتها بمنشورات المكتبات الآخرى .	Al Daihani and Abrahams (2016)
	تقييم مدى رضا المستفيدين عن مصادر المكتبة وإشراكهم في الاختيار	Al Daihani and Abrahams (2016)
	تعد البيانات المستخرجة من مواقع التواصل الاجتماعي بمثابة ثروة حقيقية	Reinhalter & Wittmann (2014)
	اتخاذ القرار وخلق قيمة مضافة	علي بن ذيب (٢٠١٧)
	التنمية المستدامة	أمل حسين (٢٠١٨)
	صناعة المعلومات	آمال بنت سعود العتيبي (٢٠١٨)
التطبيق الفعلي لتقنيات التحليلات في المكتبات	أسلوب Chernoff	Young seak (2017)
	تقنية المرئيات Visualization خلال نظم المعلومات الجغرافية GIS	Lauren Mandel (2016)

إختصاصي المكتبات والمعلومات و تحليلات البيانات

أشار (Sandhu 2015) إلى أهمية تعليم تنقيب وتحليل البيانات في المكتبات ، وذلك لتعزيز آداء المكتبات والخدمات المقدمة ، كما أشار (Reinhalter and Wittman 2014) إلى أن فرص استغلال تحليلات البيانات استحوذت على إهتمام المكتبة . وفيما يتعلق بالبحث العلمي ، يمكن لأمناء المكتبات القضاء على الفجوة الوظيفية من خلال تنفيذ المتطلبات والممارسات الجيدة بالتعاون مع منصات إدارة البيانات Data base management Platform ، ويستطيع أمناء المكتبات القيام بذلك حيث يمتلكون الخبرة في معايير البيانات والأدوات والنماذج Models التي تمكنهم من الخوض في علم البيانات بشكل عام ، وبالتالي تعاونهم مع الباحثين في خلق نماذج جديدة تدعم وتصور ما يسمى بالبيانات الضخمة فضلاً عن وضع الأدوات الخاصة بتحليل البيانات ، كما يسدي أمناء المكتبات المشورة إلى الباحثين بشأن تخطيط

وإدارة البيانات وتحديد معايير البيانات الوصفية وأساليب الحصول على البيانات وتسجيلها (علا نبيل ٢٠١٩) .

واقترح (Rani 2016) أنه لابد من التعاون بين أمناء المكتبات لانشاء مستودعات للبيانات ، ويمكن القضاء على تحديات ذلك من خلال التكيف مع التكنولوجيات الجديدة ، وبالتالي تحسين إتجاهاتها المتطورة ومساعدة الباحثين من خلال تضمين السجلات في المستودعات وقواعد البيانات حسب ترتيب الأفضلية وإهتماماتهم . كما أنهم يلعبوا دور جديد في جمع وتطوير الاحصائيات وسجلات استخدام الفهارس وقواعد البيانات بالمكتبات ، فضلاً عن إدارة البيانات البحثية ، لكن ذلك يتطلب فهم طبيعة هذه البيانات الواجب إدارتها وتحليلها والاحصائيات الواجب إنشائها من أجل البدء الفعلي وإجراء التقييم .

كما ناقش (Keil 2014) الدور الجديد لأمناء المكتبات في دعم الباحثين ، وذلك للتحكم في البيانات والاحتفاظ بها ومشاركتهم الاحصائيات ، كما ناقش القضايا الرئيسية المتعلقة بالبيانات الخام Raw Data ودور ما يسمى بالمكتبات التعليمية Educational Libraries في إيداع هذه البيانات الخام ، وذلك للحصول على نتيجة وهي " مشاركة الاحصاءات " ، وأكد على أن فكرة " خطط مشاركة البيانات الرسمية Formal Data Sharing Plan بعد تصنيفها وتحليلها تشجع الباحثين على تتبع هذه البيانات والرغبة في التعرف على نتائج تحليلها و ما يظهر من إحصاءات . وذكر دور آخر لأمناء المكتبات ألا وهو توضيح ملكية " البيانات الخام Raw Data " قبل البدء في تجميعها وتحليلها في ظل مشكلات حقوق التأليف والنشر .

وأضاف (أحمد خيري ٢٠١٨) المسميات الوظيفية المستحدثة لأمناء المكتبات في ظل التعامل مع تحليلات البيانات ، حيث عالم البيانات Data Scientist ومهندس البيانات Data Engineer حيث التعرف على القضايا الخاصة بخوارزميات البيانات وتصور وعرض البيانات ، ومعرفة تكنولوجيات قواعد البيانات مثل Hadoop ، NOSQL ، واتفقت (علا نبيل ٢٠١٩) في ذلك مع إضافة محلل البيانات الوصفية Metadata Analyst ، باحث الأعمال Business Research ، المنسق التكنولوجي Technology Coordinator حيث التكامل مع بعضهم البعض لتنفيذ التحليلات بشكل فعال وصحيح يؤخذ عليه قرار مثمر لصالح المكتبة

ومن أهم المهارات الواجب التحلي بها واكتسابها لأمناء المكتبات مهارات برمجية مثل التعامل مع منصة Hadoop، ومهارات إحصائية تحليلية للقدرة على استخلاص نتيجة من كميات البيانات وإجراء العمليات الحسابية المعقدة ، ومهارات عرض البيانات Visualization ومهارات العمل في فريق حيث أن تطبيق مثل هذه التقنيات يتطلب الخبرات المتنوعة والعمل الجماعي . (علا نبيل ، ٢٠١٩) ، واتفقت (أسماء بنت علي السلامية ٢٠١٨) في ذلك خصوصاً في التعامل مع البيانات البحثية حيث إدارتها وتحليلها عن طريق تحديد أولاً: موفر البيانات حيث إيجاد مصادر البيانات المفتوحة Open data ، وإنشاء البيانات الوصفية Metadata ، ثانياً : مقدم خدمة البيانات حيث توفير أدوات تحليل البيانات ودعم إدارة البيانات مثل خصوصية وأمن البيانات ، ثالثاً : عميل الخدمة أي المستفيد النهائي وكما يمكن للمستفيد نفسه أن ينتج خدمات جديدة بناء على نتائج تحليل البيانات . وبناء على ذلك قدم (يوسف على الشيخ ٢٠١٨) توصية بضرورة ابتعاث أخصائي المكتبة إلى دورات متقدمة في تحليل البيانات لتطوير قدراتهم . وللوصول إلى هذه الدورات التدريبية رصدت (علا نبيل ٢٠١٩) أهم المواقع المقدمة لهذه الدورات في مجال تحليلات البيانات مثل موقع Simplilearn - Coursera ومن أهم الدورات Big Data Analysis .

ويلخص الجدول التالي الموضوعات المتعلقة بدور إختصاصي المكتبات والمعلومات حسب أدبيات التخصص كالتالي:-

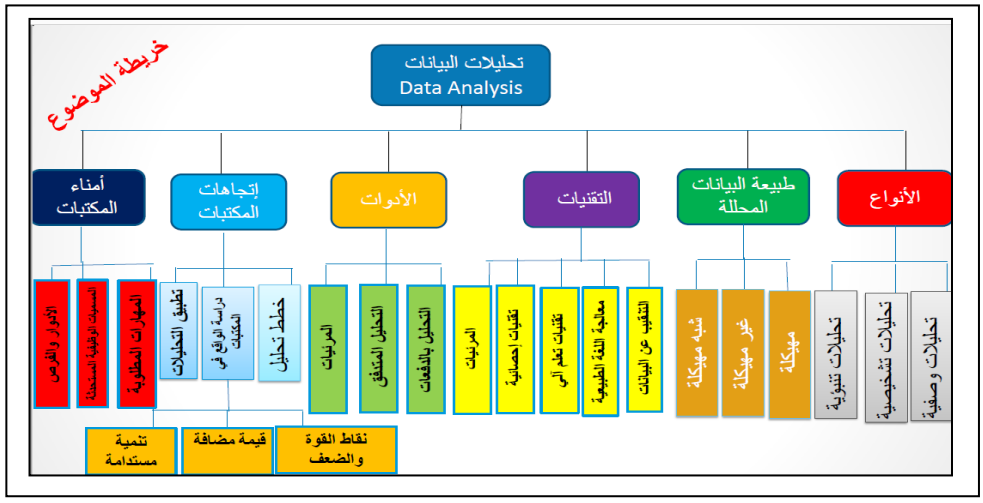
الموضوع	الدلالة	الدراسة
فهم و استيعاب قدرات وإمكانيات تحليلات البيانات	استخدام سجلات الاستخدام Usage Record للمستفيدين بشكل فعال	Reinhanlter and Wittmann(2014)
إستخدام أدوات تحليلات البيانات	استخدام أدوات إحصائية وتقنيات التنقيب عن البيانات في المكتبات	Sandhu (2015)
	جمع وتطوير وحفظ الاحصائيات	Gordon – Murane (2012)
	التدريب على استخدام التقنيات المناسبة وفقاً لأهداف الاحصائيات المرجو تحقيقها	Witt (2012)
إيداع البيانات المعقدة	تحويل البيانات المهيكلة إلى مستودعات معلومات مرئية غير مهيكلة	Keil (2014)
مشاركة نتائج تحليلات البيانات	تتبع المعلومات الناتجة عن تحليل البيانات والتعريف بها والتعريف بمشاكل الملكية الفكرية لهذه البيانات الخام التي تتم معالجتها	Keil(2014)
دعم مستخدمي البيانات المعقدة	تقديم دراسات حول الاحصائيات الناتجة عن التحليل	Keil(2014)
خلق برامج توعوية حول تحليل البيانات	برامج تدريبية	Rani(2016)
فهم التحديات المتعلقة بتحليل البيانات	حصرها ووضع خطط لإدارتها	Daniel (2015)
المسميات الوظيفية المستحدثة	عالم ومهندس البيانات ومحلل البيانات الوصفية .. الخ	أحمد خيري(٢٠١٨) ، علا نبيل (٢٠١٩)
مهارات أمناء المكتبات	تحليلة ، برمجية ، احصائية	علا نبيل (٢٠١٩) أسماء بنت علي السلامية (٢٠١٨)
دورات تدريب أمناء المكتبات	ضرورة إبتعاثهم إلى دورات متقدمة في تحليلات البيانات	يوسف علي الشيخ (٢٠١٨)
	رصد المواقع المقدمة للدورات	علا نبيل (٢٠١٩)

وبعد عرض المحاور الرئيسية والفرعية لموضوع " تحليلات البيانات في المكتبات ومراكز المعلومات " ، إتجهت الباحثة لحساب النسبة المئوية لمدى إستحواذ كل محور من محاور الموضوع على الإنتاج الفكري العربي والإنجليزي ، ويتضح ذلك في الجدول التالي :-

المحور	عدد المصادر	النسبة المئوية
طبيعة البيانات الواجب تحليلها	٢٩	%٤٦,٧٧
تقنيات تحليلات البيانات	١٥	%٢٤,١٩
أدوات تحليلات البيانات	١٩	%٣٠,٦
إتجاهات المكتبات ومراكز المعلومات في تحليلات البيانات	٢٢	%٣٥
إختصاصي المكتبات والمعلومات و تحليلات البيانات	١٥	%٢٤,١٩

ويتضح من الجدول السابق أن المصادر المعتمد عليها في معالجة المحور الأول ألا وهو طبيعة البيانات الواجب تحليلها حازت على المركز الأول بنسبة وهي ٤٦,٧٧% وذلك يرجع إلى ان معظم المصادر العربية والإنجليزية كانت تتناول هذا الجزء النظري ، ويليها المركز الثاني في المحور الخاص بإتجاهات المكتبات ومراكز المعلومات في تحليلات البيانات بنسبة ٣٥% ، وذلك بسبب تنوع الإنتاج الفكري العربي والإنجليزي في تناول الموضوع وأضاحت الباحثة ذلك في المحاور الفرعية للجدول الخاص بهذا المحور ، وجاء في المركز الثالث محور تقنيات تحليلات البيانات بنسبة ٢٤,١٩% وذلك يرجع إلى أن هذا المحور تم تناوله في الإنتاج الفكري عربي وإنجليزي من زاويتين وهما زاوية التخصص " مكتبات ومعلومات " وزاوية تقنية برمجية " علوم الحاسب والخوارزميات " ، وجاء في المركز الأخير بالتساوي أدوات البيانات الضخمة وإختصاصي المكتبات والمعلومات بنسبة ٣٠,٦% وذلك يرجع إلى أن المسمى الأكثر شيوعاً في الجزء التقني تمثل في التقنيات وليس الأدوات ، ولكن بعد البحث بشكل أكثر عمقاً تمكنت الباحثة من تقسيم هذه الأدوات وفقاً للتقنيات الرئيسية بالاعتماد على هذه المصادر ، وبالنسبة لإختصاصي المكتبات فتم ترجمه سبب الفجوات هذه في شكل إتجاهات بحثية مستقبلية .

وبعد الإنتهاء من المرحلة السابقة ، تمكنت الباحثة من رسم خريطة معرفية لموضوع تحليلات البيانات بكل جوانبه وإتجاهاته ، ويتضح ذلك في الشكل التالي :-



شكل رقم (٣) الخريطة النهائية لموضوع تحليلات البيانات

نتائج تحليل الإنتاج الفكرى حول موضوع تحليلات البيانات

بعد عرض ما سبق من أدبيات الموضوع في عدة محاور ، لوحظ العرض الوافي لطبيعة البيانات الواجب تحليلها بأدبيات الموضوع العربية والإنجليزية وكذلك أهم السمات المميزة لها وتعريفها من وجهات نظر مختلفة ، وبالنسبة للمحور الخاص بالتقنيات الخاصة بتحليلات البيانات تم معالجتها في الإنتاج الفكري الأجنبي بشكل منفصل عن الأدوات ، حيث وضع التقنية بمثابة الموضوع الرئيسي ويندرج تحت كل تقنية عدة أدوات يتم العمل معها لإجراء عملية التحليل . كما لوحظ تفوق الإنتاج الفكري العربي حول رصد تقنيات المرئيات ورصد أهم البرمجيات مثل Timeline - Gephi - Tableau .. الخ .

وبالنسبة للجزء التقني في موضوع تحليلات البيانات ، تمت معالجتة في الإنتاج الفكري العربي بكليات التجارة والحاسبات والمعلومات ، وفي الإنتاج الفكري الأجنبي في مدارس التعلم الآلي وعلوم الحاسوب ، ولكن في تناول المحتوى تفوق الإنتاج الفكري الأجنبي في الشرح الوافي لكل تقنية ولكل أداة برمجية حول الموضوع ، حيث وجدت دارسات حالة على سبيل المثال دراسة حالة لتقنية S3 حيث دراسة طبيعتها وتحليلها البرمجي . في حين تناول الإنتاج الفكري التقنية ولكن ليست بشكل مستقل أي التعرف على طبيعتها خلال دراسة طبقتها على مؤسسة ما وهكذا.

كما لوحظ معالجة الاتجاهات المختلفة لتحليلات البيانات في المكتبات ومراكز المعلومات بشكل واف في الجزء الخاص بدراسة الواقع ونقاط القوة والضعف حول تطبيق عملية التحليل ووضع الخطط والمشروعات لتطبيق تحليلات البيانات ، ولكن وجد قصور في الجزء الخاص بالخطط حيث تم تناول الأطر الرئيسية وأهم التقنيات الواجب استخدامها ويتم الاطراد في عرض الخطوات الأولية والنهائية المتمثلة في إتخاذ القرار ، لكن الجزء الخاص بتطبيق التقنية وإيضاح معايير تطبيق هذه التقنية لم يتم تناولة في الإنتاج الفكري بتخصص المكتبات بشكل واف حتى وان تم تطبيقها بالفعل ، باستثناء تقنية واحدة من تقنيات تحليل البيانات وهي المرئيات Visualization باتباع أسلوب Chernoff وتم ذكرها بالتفصيل في الانتاج الفكري الإنجليزي . وتفوق الإنتاج الفكري الإنجليزي أيضاً في التطبيق الفعلي لتحليلات البيانات ورصد ما أنتجته من فوائد مثمرة للمكتبة دون الاقتصار على دراسة الواقع .

وبالنسبة لدور أمناء المكتبات تم تناوله في الانتاج الفكري العربي بشكل أكثر تفصيلا في الجزء المختص بأهم المهارات الواجب توافرها والمواقع المقدمة لمثل هذه الدورات ، أما في الإنتاج الفكري الإنجليزي كان التركيز على الأدوار الأساسية المتعلقة بإيداع البيانات المعقدة ومشاركة نتائج التحليل ودعم الباحثين .. الخ ، ولم يتطرق أحد إلى معالجة القوانين والتشريعات الخاصة بتنظيم وتحليل البيانات غير المهيكلة بواسطة أمناء المكتبات داخل مؤسسات المكتبات ومراكز المعلومات .

ووجد قصور من وجهة نظر الباحثة في نقاط هامة مثل الجانب الاقتصادي الناتج عن تحليلات البيانات في المكتباب ، تطبيق تقنيات التحليلات بالفعل في مكتباتنا العربية ، الخطط الفعلية لتدريب أمناء المكتبات، أوجه المقارنة بين تحليل المحتوى المهيكل وغير المهيكل بالمكتبات ومؤسسات المعلومات ، ودراسة التحدي الخاص بالتكلفة حول تطبيق هذه التقنيات بالمكتبة و أيضاً الجانب الخاص بمقررات تحليلات البيانات في أقسام المكتبات والمعلومات ، وسوف يتم ترجمة هذا القصور في شكل إتجاهات بحثية توصي الباحثة بتناولها مستقبلياً .

وقبل تناول الاتجاهات المستقبلية ذلك يمكن تلخيص الإتجاهات البحثية التي **تم تناولها بالفعل** في الإنتاج الفكري العربي والأجنبي والتى تم عرضها فى هذه المراجعة فى الجدول التالى

موضوعات بحثية ذات علاقة بالمكتبات ومراكز المعلومات	موضوعات بحثية ذات علاقة بإختصاصي المكتبات والمعلومات
تحليل خطط ومشروعات تحليلات بيانات قائمة بالفعل بالمكتبات الأجنبية	تحليل سجلات الاستخدام User Record للمستفيدين
دراسة واقع تطبيق تحليلات البيانات في المكتبات العربية والأجنبية	إستخدام أدوات إحصائية في المكتبات
تطبيق تقنيات تحليل محتوى منشورات وسائل التواصل الإجتماعي بالمكتبات الأجنبية	إيداع البيانات في مستودعات معلومات مرئية
دور المكتبات في تحليلات نوع هام من البيانات وهو البيانات البحثية	مشاركة نتائج التحليلات
تطبيق تقنيات المرئيات بالمكتبات الأجنبية	المسميات الوظيفية المستحدثة لأمناء المكتبات والمهارات اللازم توافرها للتعامل مع التحليلات
	خلق برامج توعوية للمستفيدين حول التحليلات ، و حضور دورات تدريبية لتمنيه المهارات

الإتجاهات البحثية المستقبلية

بالإعتماد على الجدول السابق للموضوعات البحثية التي تم تناولها بالفعل ، يمكن إقتراح عدة إتجاهات بحثية مستقبلية بناء على هذه الموضوعات وبناء على تحليل محتوى الإنتاج الفكري المندرج تحت هذه الموضوعات ، ومن هذه الاتجاهات :-

أوجه القصور	الدراسات المقترحة
معالجة الجانب الاقتصادي نتيجة تحليلات البيانات في المكتبات	دراسة مدى تأثير تحليلات البيانات بشكل منهجي على القيمة الإقتصادية في المكتبات الأكاديمية .
التطبيق الفعلي لتحليلات البيانات في المكتبات ومؤسسات المعلومات العربية	١. دراسة تحليلية لأحد الأدوات المستخدمة في تحليل المواد الاعلامية الاجتماعية لأحد المكتبات العربية للإعلام بمدى الاستفادة منها . ٢. دراسة تطبيقية لتحليلات البيانات باستخدام Hadoop في المكتبة القومية الزراعية في مصر على غرار المكتبة القومية الزراعية بالولايات المتحدة الأمريكية . ٣. تحليلات البيانات بالمكتبة القومية الزراعية في مصر : دراسة للواقع وتصور للمستقبل ٤. دراسة تطبيقية حول تجميع وتصنيف البيانات المتعلقة بتحليلات البيانات كموضوع وإتاحتها للمستفيدين عبر الوجه أى الارشاد الرقمي على غرار مكتبة جامعة هونغ كونغ في الصين والمطلق عليها Digital Scholarship Services .

أوجه القصور	الدراسات المقترحة
	٥. تطبيق أسلوب Chernoff لتحليل البيانات المتعلقة بأوعية المعلومات داخل المكتبة ، حيث كل وعاء يأخذ معايير وصف الأسلوب حيث وصف البيانات التالية أهم الموضوعات الفرعية ، عدد مرات الاستعارة ، عدد المؤلفين ، الشكل المادي ، عدد الصفحات .. الخ . ٦. مشروع بحثي حول تنفيذ مراكز البحوث في الجامعات بتحليل البيانات التي تمتلكها الجامعات والمكتبات الجامعية لتحقيق فوائد أمنية واقتصادية واجتماعية مع مراعاة خصوصياتها.
وضع خطط فعلية لإختصاصي المكتبات والمعلومات	١- دراسة لوضع تصور مقترح للمبادئ التوجيهية لأمناء المكتبات بشأن منهجية استخدام التحليلات بالمكتبات بمساعدة بعض الجهات المختصة . ٢- دراسة استكشافية حول خطة تدريب أمناء المكتبات المتبعة للتعامل مع تحليلات البيانات في بعض المكتبات الأجنبية وأخذها كدليل لمكتباتنا العربية .
أوجة المقارنة بين تحليل المحتوى المهيكل وغير الميهكل	١- دراسة وصفية مقارنة لعملية تصنيف البيانات غير المهيكلة بالمكتبات ومقارنتها بالآخرى المهيكلة . ٢- دراسة مقارنة بين تكشيف محتوى مصادر المعلومات المهيكل والمحتوى الآخر غير المهيكل مثل التغريدات والمنشورات لموقع المكتبة .
معالجة تحديات تحليلات البيانت	دراسة التحدي المتعلق بالتكلفة الخاصة باقتناء برمجيات تحليلات البيانات بالمكتبة حيث يعتبر عائق أمام المكتبات العربية ، ووضع حلول مقترحة .
مقررات تحليلات البيانات بأقسام المكتبات والمعلومات	١- دراسة استكشافية لمدى تدريس تحليل البيانات أو ما يرتبط به من تقنيات في أقسام المكتبات والمعلومات بالمكتبات الأجنبية ووضع مقترح لبلورته في مقرراتنا العربية . ٢- دراسة تجريبية لتدريس مقرر خاص بتحليلات البيانات ببرنامج الماجستير أو الدكتوراة بأقسام المكتبات والمعلومات في مصر .

قائمة المصادر التى تم عرضها فى المراجعة العلمية :

تم توثيق المصادر بأحد برمجيات التوثيق الإلكتروني طبقا لقواعد التوثيق MLA 8 ورابطة http://www.easybib.com/mla8-format/website-citation

أولا : المصادر باللغة العربية

- أحمد ، ثابت محمد." البيانات الضخمة ورهانات الأدوار المتغيرة لأخصائي المعلومات : مقارنة فكرية ورصد تحليلي للكفايات والمواصفات المستقبلية في ضوء تقنيات الويب ٠,٢ " . المؤتمر والمعرض السنوي الرابع والعشرين لجمعية المكتبات المتخصصة فرع الخليج العربي . ٢٠١٨ ، ص ص ١-٢٧.

- أحمد ،فايز أحمد." نظم إدارة قواعد البيانات الضخمة : دراسة حالة لنظام أباتشي هادوب Hadoop Apache ، الاتحاد العربي للمكتبات والمعلومات ، ٢٠١٩ ، ص ص ١٢١-١٨٤ .
- الأكذبي ،علي بن ذيب ." تحويل البيانات الضخمة إلى قيمة مضافة " ، مجلة مكتبة الملك فهد الوطنية ، ٢٠١٧ ، ص ص ٨٢-١٠٢ .
- بو عناقة ، سعاد . " البيانات الضخمة في قطاع المكتبات : نقاط القوة والضعف والفرص والتهديدات " ، المؤتمر الرابع والعشرون : البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفي ، ٢٠١٨ ، ص ص ١- ١٧ .
- حمدان ، منال سعيد ." البيانات الضخمة في المكتبات الأكاديمية في سلطنة عمان " ، المؤتمر الرابع والعشرون : البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفي ، ٢٠١٨ ، ص ص ١- ٢٠
- خالد عتيق ، عبدالله بن سالم ." البيانات الضخمة في مكتبات جامعة السطان قابيوس : واقعها ومستوى الافاردة منها ، المؤتمر الرابع والعشرون : البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفي ، ٢٠١٨ ، ص ص ١- ٢٨.
- خلود بنت خالد ، هاجر سليمان ." واقع استخدام البيانات الضخمة في هيئة الوثائق والمحفوظات الوطنية : دراسة استطلاعية " ، المؤتمر الرابع والعشرون : البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفي ، ٢٠١٨ ، ص ص ١- ١٧
- خيري ، أحمد عبدالله ." دور مراكز المعلومات في إدارة البيانات الضخمة : مؤسسات الرعاية الصحية في مصر نموذجاً " ، رسالة دكتوراة ،جامعة سوهاج ، كلية الآداب ، قسم المكتبات والمعلومات ، ٢٠١٨ ، ص ص ١-٢٨٨ .
- راشد ، أسماء بنت سعيد." البيانات الإحصائية والبيانات الضخمة : واقع إنتاج وإستثمار البيانات الاحصائية في المركز الوطني للإحصاء والمعلومات بسلطنة عمان " ، المؤتمر الرابع والعشرون : البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفي ، ٢٠١٨ ، ص ص ١-٦٥ .
- السلامية ، أسماء بنت علي ." البحث العلمي في عصر البيانات الضخمة : الأدوار والفرص لاختصاصي المعلومات " ، المؤتمر الرابع والعشرون : البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفي ، ٢٠١٨ ، ص ص ١-١٩ .
- عبد القادر ، أمل حسين." إسهام البيانات والمعلومات في التنمية المستدامة : رؤية مستقبلية " ، المؤتمر الرابع والعشرون : البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفي ، ٢٠١٨ ، ص ص ١-١٩ .
- العتيبي أمال بن سعد. "البيانات الضخمة وصناعة المعلومات في المكتبات ومراكز المعلومات : مكتبة الملك فهد الوطنية نموذجا" .المؤتمر والمعرض السنوي الرابع والعشرين لجمعية المكتبات المتخصصة فرع الخليج العربي . ٢٠١٨،ص ص ١-١٩.
- فتوح ، عمرو حسن. "تحليلات البيانات الضخمة ودورها في دعم إتخاذ القرار في المكتبات ". مجلة المكتبات والمعلومات العربية ، أكتوبر٢٠١٧ ،ص ص ١-٣٢ .

- القصاص ، علا نبيل ." مشروعات البيانات الرقمية الضخمة في المكتبات : دراسة استكشافية " ، رسالة ماجستير ، جامعة المنوفية ، كلية الآداب ، قسم المكتبات والمعلومات، ٢٠١٩ ، ص ص ١-٢٠١.
- مصطفى ، يوسف علي الشيخ ." التخطيط للأدوار الأساسية لأخصائي المكتبات في التعامل مع البيانات الضخمة بالإشارة إلى مركز الفيصل الثقافي – السودان " ، المؤتمر الرابع والعشرون : البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفي ، ٢٠١٨ ، ص ص ١-٢١.
- النشرتي، مؤمن. "نحو التكامل المعرفي من واقع توظيف الأنطولوجيات في أطار التنقيب عن البيانات: دراسة تحليلية Cybrarians Journal ".،٢٠١٤، ص ص ١-٣٤.
- نجيب ، محي الدين ." دور وأهمية المكتبات الجامعية في إدارة البيانات الضخمة : دراسة حالة المكتبة المركزية جامعة الخرطوم " ، المؤتمر الرابع والعشرون : البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفي ، ٢٠١٨ ، ص ص ١-٢٣ .
- وزارة تكنولوجيا المعلومات والاتصالات بقطر ." "البيانات الضخمة :تحقيق التوازن بين المزايا والمخاطر"، ٢٠١٤، ص ص ١-١٣ .

ثانيا : المصادر العربية بالدوريات التقنية حول الموضوع

- Mohammed Malhat and etal .” Anovel scalable and effective partitioning for big data reduction “ , Menofia university , IJCI,vol6,nol1,2019,pp9-19.
- Elham abelaziz .” Data envolvement analysis : A technique for measuring efficiency “ , Al –azhar university , faculty of commerece ,2012,pp23-46.
- Sally ali,hamdy mousa , m.hussien .” A review of open information extraction tecgniques “ , IJCI,vol6,2019,pp20-28

ثالثا : المصادر باللغة الإنجليزية

- Akoka, J., Comyn-wattiau, I., & Laou, N. "Research on Big Data – A systematic mapping study". Computer science , 2017,pp105-115.
- Al-Daihani, S., & Abrahams, A.” A text mining analysis of academic libraries’ tweets”. The Journal of Academic Librarianship, 42,2016, 135–143
- Andreu-Perez, J., Poon, C. C., Merrifield, R. D., Wong, S. T., Yang, G.-Z..” Big Data for health. IEEE Journal of Biomedical and Health Informatics”, 19(4),2015, 1193–1208. DOI: 10.1109/JBHI.2015.2450362
- Archenaa, J., & Mary Anita, E. A. “A survey of Big Data analytics in healthcare and government”. Procedia computer science ,2015,pp408-413.
- Begoli, E. & Horey, J. “ Design principles for effective knowledge discovery from big data.” In Joint ICSA and ECSA,2012, 215–218. Available from: http://www.bdva.eu/sites/default/files/Design%20Principles%20for%20
- Beyer, M. A., & Laney, D. “ The importance of ‘Big Data’: A definition. Stamford, CT: Gartner,2012.

- Bizer, C., Boncz, P., Brodie, M. L., & Erling, O. " The meaningful use of Big Data: Four perspectives." SIGMOD, 40(4,)2011, 56–60
- Borkar, V., Carey, M. J., Li, C. " Inside "Big Data management": Ogres, onions, or parfaits? In Proceeding of EDBT/ICDT joint conference. Berlin: ACM,2012.
- Bostock, M., Ogievetsky, V., & Heer, J." D3 data-driven documents." IEEE Transactions on Visualization and Computer Graphics, 17(12),2011, pp 2301–2309
- Bu, Y., Brokar, V., Carey, M. J., Rosen, J., Polyzotis, N., Condie, T., … Ramakrishnan R." Scaling datalog for machine learning on Big Data. Computer research repository (CoRR) ", 2012, pp. 1–14,. Cornell University Library. DOI: http://arxiv.org/pdf/1203.0160v2.pdf
- Chen, H., Chiang, R. H. L., & Storey, V. C. " Business intelligence and analytics: From Big Data to big impact. MISQ", 36(4),2012 , 1165–1188
- Cuzzocrea, A., Song, I. Y., & Davis, K." Analytics over large-scale multidimensional data: The Big Data revolution! "Proceedings of the 14th international workshop on Data Warehousing and OLAP, New York, NY: ACM,2014,pp101-103.
- Daniel, B." Big Data and analytics in higher education: Opportunities and challenges". British Journal of Educational Technology, 46(5),2015, 904–920. DOI: 10.1111/bjet.12230 Effective%20Knowledge%20Discovery%20from%20Big%20Data.pdf
- Fan, W., & Bifet, A." Mining Big Data: Current status, and forecast to the future". SIGKDD Explorations, 14(2),2012,pp 1–5.
- Franceschini, M. " How to maximize the value of Big Data with the open source SpagoBI suite through a comprehensive approach" . Proceeding of the VLDB Endowment, 2013, 1170–1171.
- Gantz, J., & Reinsel, D. " Extracting value from chaos". IDC iView,2011, 1–12.
- Gordon-Murnane, L. " Big Data: A big opportunity for librarians". Online, 36(5),2012, 30–34.
- Han, J., Kamber, M., & Pei, J. "Data mining: Concepts and techniques "(3rd ed.). Waltham, MA: Elsevie,2011.
- Hashem, I. A. T., Yaqoob, I., Badrul Anuar, N., Mokhtar, S., Gani, A., & Khan, S. U. "The rise of "Big Data" on cloud computing: Review and open research issues". Information Systems, 47, 2011,pp 98–115.
- Heidorn, P. B. " The emerging role of libraries in data curation and e-science". Journal of Library Administration, 51(7–8),2011, 662–672.

- Herodotou, H., Lim, H., & Luo, G. " Starfish: A self-tuning system for Big Data analytics". In Proceeding of the 5th biennial conference on innovative data systems research (CIDR 11),2011, pp. 261–272.
- Jody condit Fagan ." the effects of academic libraries : resources , expenditure and service decisions library use : An Analysis of ACRL and NCES data , James maison university , school of strategic leadership , 2014 , pp1-246 .
- Keil, D. "Research data needs from academic libraries: The perspective of a faculty researcher". Journal of Library Administration, 54(3),2014, 233–240.
- Khan, S., Liu, X., Shakil, K. A., & Alam, M." A survey on scholarly data: From big data perspective". Information Processing & Management, 53(4), 2017,pp 923–944.
- Khurshid ahmed , zhengming , muhammed rafi ." analysis of academic librarians competencies , and skills for implementation of big data analytics in libraries , emerald , vol 5 , n2 , 2019, pp1-16 .
- Larkou, G., Mintzis, M., Andreou, P. G., Konstantinidis, A., & Zeinalipour-yazti, D. " Managing Big Data experiments on smartphones." Distributed and Parallel Databases, 34(1),2014, 33–64. DOI: http://doi.org/10.1007/ s10619-014-7158-6
- Louren mandel . " visualizing the library as place " , USA , university of rhode island , 2016 , pp1-10.
- Madden, S. " From databases to Big Data". IEEE Internet Computing, 2012,16(3), 4–6.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. "Big Data: The next frontier for innovation, competition and productivity". New York, NY: McKinsey Global Institute,2011.
- Neumeyer, L., Robbins, B., Nair, A., & Kesari A. "S4: Distributed stream computing platform". In Proceeding of the 2010 international conference on data mining workshops (ICDMW). IEEE,2010.
- Rani, B. R. " Big Data and Academic Libraries.", In International conference on Big Data and knowledge discovery. Indian Statistical Institute,2016.
- Reinhalter, L., & Wittmann, R. J." The library: Big Data's boomtown", The Serials Librarian, 67(4),2014, 363–372. DOI: 10.1080/0361526X.2014.915605
- Rodrيguez-Mazahua, L., Rodrيguez-Enrيquez, C. A., Sلnchez-Cervantes, J. L., Cervantes, J., Garcيa-Alcaraz, J. L., & Alor-Hernلndez, G. " A general perspective of Big Data: Applications, tools, challenges and trends". The journal of supercomputing , 2016,pp3113-3073.

- S,Young , C.louise .” big data analysis of public library operations and services by using the Chernoff face method “ , journal of Documentation , vol 73 , 2017 , pp1-15.
- Sandhu, G.” Re-envisioning library and information services in the wake of emerging trends and technologies”. The 4th international symposium on emerging trends and technologies in libraries and information services, Noida, India ,2015, pp. 153–160
- Siguenza-Guzman, L., Saquicela, V., Avila-Ordٌٓٓez, E., Vandewalle, J., & Cattrysse, D. “ Literature review of data mining applications in academic libraries”. The Journal of Academic Librarianship,2015, 41(4), 499–510.
- Slavakis, K., Giannakis, G. B., & Mateos, G. “ Modeling and optimization for Big Data analytics”. IEEE Signal Processing Magazine,2014, 31(5),pp 18–31.
- Stoica, I. “ Conquering Big Data with spark and BDAS”. In Proceeding of the ACM international conference on measurement and modeling of computer systems,2014.
- Suthaharan, S. “ Big Data classification: Problems and challenges in network intrusion prediction with machine learning “ , 2014.pp1-9.
- Ward, J. S., & Barker, A. .” Undefined by data: A survey of Big Data definitions”. 2013 Available from: http://arxiv.org/pdf/1309.5821v1.pdf .
- Witt, M. “ Co-designing, co-developing, and co-implementing an institutional repository service”. Journal of Library Administration, 52(2),2012, 172–188
- Zhifeng, X., & Yang, X. “Security and privacy in cloud computing.” IEEE Communications Surveys and Tutorials, 15(2),2013, 843–859.

الاستشهادات

- Hart , Harold (1968) Critical reviews the editors point of view . J. of chemical documentation , vol 8 , no 4 (Nov.) p 241 – 244 .: فى
- قاسم ، حشمت محمد ." المراجعات العلمية ودورها في تحويل المعلومات إلى معرفة : إعداد المراجعات العلمية ونشرها الجزء الثاني " ، دراسات عربية فى المكتبات والمعلومات
- Virgo , Julie (1971)The review article :it’s characteristics and problems , the library quarterly , vol.41 , 1971,p281.
- مجمع اللغة العربية (٢٠١١) . " المعجم الوسيط " ، القاهرة : مكتبة الشروق الدولية ،ط٥ ، ص ٨١
- وزارة تكنولوجيا المعلومات والاتصالات بقطر (٢٠١٤) البيانات الضخمة : تحقيق التوازن بين المزايا والمخاطر . قطر : الوزارة ، ص ١٠ .