إدارة البيانات البحثية في قاعدة بيانات ISI Web of Science : دراسة تحليلية باستخدام أسلوب التحليل الببليومتري والتحليل المرئي

أ.د. سهير عبد الباسط عيد
قسم علوم المعلومات – كلية الآداب
جامعة بني سويف

المستخلص:

تهدف هذه الدراسة إلى رصد وتحليل الإنتاج الفكري في مجال "إدارة البيانات البحثية" المكشف في قاعدة بيانات ISI Web of Science، باستخدام الأساليب الببليومترية لدراسة التوزيعات المختلفة لهذا الإنتاج الفكري، و باستخدام برنامج VOSviewer لتصوير البيانات وتحليلها بغية التحقق من التواجد المشترك للكلمات المفتاحية في وحدات الإنتاج الفكري المدروس. وقد تبين من الدراسة أن عام 2019م شهد أعلى معدل نشر بواقع 88 عملا، وأن مقالات الدوريات هي أكثر أشكال مصادر المعلومات ووصل عددها إلى 379 مقالة، وجاء مجال المكتبات والمعلومات في المرتبة الأولى من بين المجالات الموضوعية التي نشرت في إدارة البيانات البحثية بواقع 310 عملاً، ومن حيث أماكن النشر جاءت الولايات المتحدة الأمريكية في المقدمة بعدد 138 عملا، و تصدرت اللغة الإنجليزية بنسبة 89.56 %، وجاء المؤلف البرتغالي João Rocha da Silva. أعلى المؤلفين إنتاجية بعدد 15 عملاً، وجاءت المقالة بعنوان Panorama: A Targeted Proteomics Knowledge Base المنشورة عام 2014م أعلى المقالات المستشهد بها، واعتمادًا على شبكات الكلمات المفتاحية التي توضح المصطلحات الأكثر ترددًا بالإضافة إلى المصطلح الأساس إدارة البيانات البحثية جاءت مصطلحات : المكتبات، والبيانات البحثية، ومعالجة البيانات، ومبادئ FAIR للبيانات كموضوعات ذات علاقة بمصطلح إدارة البيانات البحثية.

الكلمات المفتاحية: إدارة البيانات البحثية ؛ التحليل الببليومتري ؛ التحليل المرئي؛ شبكة الكلمات المفتاحية .

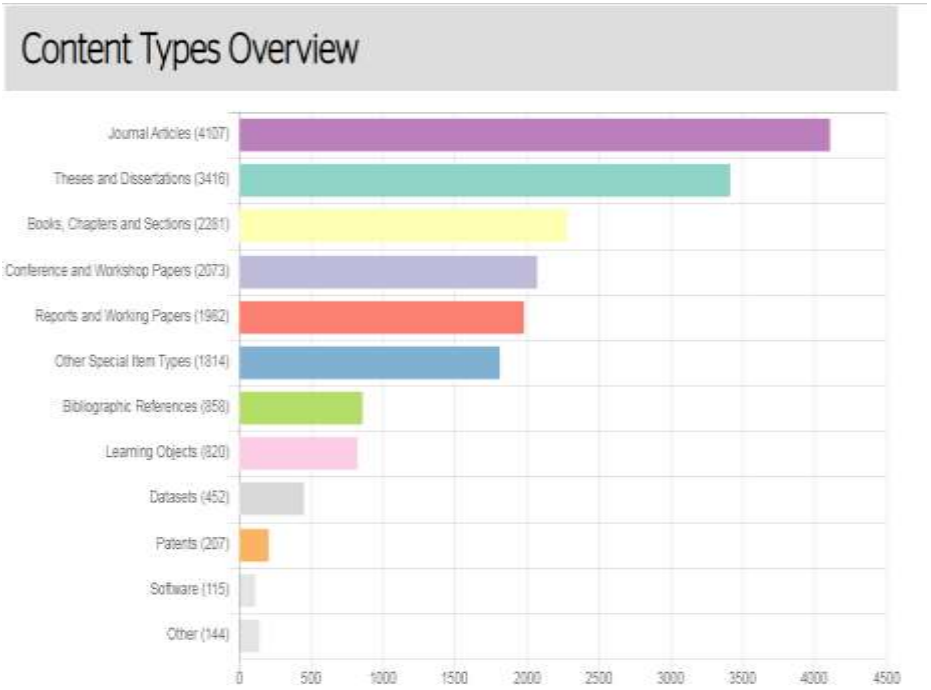
تاريخ استلام البحث: 2022/5/17 **تاريخ إجازة البحث: 2022/5/25**

1) المقدمة المنهجية:

1/1) تمهيد:

هناك تحول كبير، في المجتمع العلمي اليوم، نحو العلم المفتوح؛ الذي يعطي الأولوية للتعاون والوصول المفتوح للبيانات ومخرجات البحث العلمي. وأصبحت إدارة البيانات البحثية ومشاركتها جزءًا مهمًا بسبب أهميتها للباحثين في تطوير العلم؛ وذلك من خلال إعادة استخدام البيانات التي أنتجها باحثون آخرون.

ونلاحظ مؤخرًا أن الهيئات الممولة للبحوث تطلب من الباحثين إرفاق خطة لإدارة بيانات البحث مع مقترحاتهم، فعلى سبيل المثال تطلب المؤسسة القومية للعلوم National Science Foundation (NSF) من الباحثين أن تضم مقترحاتهم البحثية المقدمة طلبًا للدعم، وثيقة تكميلية لا تزيد عن صفحتين بعنوان "خطة إدارة البيانات". تصف هذه الوثيقة التكميلية كيف سيتوافق المقترح البحثي مع سياسة المؤسسة NSF بشأن نشر وتبادل نتائج البحث(1). كما أن الباحثين في كثير من الأحيان يستعينون بأمناء المكتبات أو الأرشيفيين لكتابة خطط إدارة البيانات البحثية، وعلى الجانب الآخر تضع مستودعات البيانات سياسات لاتخاذ قرارات بشأن استيعاب البيانات وتنظيمها ونشرها(2). ومن الملاحظ أيضًا أن أدلة مستودعات الوصول الحر العالمية، اعتمدت مستودعات البيانات البحثية أحد أنواع المحتوى الذي تكشفه وتضمه إليها، فعلى سبيل المثال ضم دليل مستودعات الوصول الحر Directory of Open Access Repositories (DOAR) مستودعات البيانات البحثية (أو الأولية) ضمن انواع المحتوى، ووصل عددها على مستوى العالم وفقًا لهذا الدليل 452 دليل بنسبة 2.47% من إجمالي المحتوى، كما يتضح من الشكل التالي(3):



شكل رقم (1) توزيع مستودعات الوصول الحر المدرجة في دليل (DOAR) تبعًا لنوع المحتوى

وبالطبع نشـــرت آلاف الأوراق العلمية والبحوث في موضـــوع البيانات البحثية على مســـتوي العـالم من حيـث أنواعهـا، وأهميتهـا في مجـال البحـث العلمي، وخطط إدارتهـا على المسـتوى المؤسـسـي، ومسـتودعاتها، ومعايير تنظيمها، ومشـاركتها، وحفظها الدائم، ومدى وعي الباحثين بها، وتتزايد هذه الدراســـات عامًا بعد عام. ورغم ذلك فإن موضـــوع إدارة البيانات البحثيـة يعـد من أهم الموضـــوعـات المتعلقـة بطوفـان البيـانـات البحثيـة؛ نظرًا لأن البيـانـات البحثية المدارة بشـكل جيد تُسـهِّل عمليات التخزين والاسـترجاع والمشـاركة وإعادة الاسـتخدام والأمن[4]

ويعد مجال إدارة البيانات البحثية مجالًا متعدد التخصـــصـــات، إذ أن مجموعات البيـانـات ينتجهـا بـاحثون من مختلف المجـالات. وبصـــفـة عـامـة يتم دراســـة التعـاون بين التخصـــصـــات المختلفـة وفقًـا لمخرجـات البحـث في مجـال مـا بطريقتين، الأولى: تعتمـد على

الأساليب الإحصائية وتحليل الشبكة الاجتماعية لتصوير أو تخطيط علاقة التواجد المشترك للكلمات المفتاحية co-occurrence، والاستشهادات، وغيرها من البيانات الببليومترية؛ ومن ثم تصنيف التخصصات والكشف عن العلاقات بينها. والثانية: يتم تحديد الهيكل متعدد التخصصات للمجال البحثي من خلال تحليل الاستشهادات في الدوريات لاكتشاف التعاون بين التخصصات الرئيسية وعلاقاتها بمرور الوقت، فمن خلال مراجع المقالات رسم الباحثون خريطة للتخصصات المتعددة وتقييمها[5].

2/1) مشكلة الدراسة وتساؤلاتها:

أصبح موضوع "إدارة البيانات البحثية" من الموضوعات الحيوية في السنوات الأخيرة، وزاد حجم الإنتاج الفكري المنشور والمكشف في قواعد البيانات العالمية، وبالتالي تظهر الحاجة ملحة لدراسة الخصائص الببليومترية لهذا الإنتاج الفكري، والتعرف على المجالات ذات العلاقة به؟ وتتلخص تساؤلات الدراسة في:

1- ما التوزيع الزمني للدراسات المكشفة والمدرجة في قاعدة البيانات WoS في موضوع إدارة البيانات البحثية؟
2- ما أشكال الدراسات المدرجة في قاعدة البيانات البيانات WoS في موضوع إدارة البيانات البحثية؟
3- ما المجالات الموضوعية التي نشر فيها أعمال في "إدارة البيانات البحثية"؟
4- ما التوزيع اللغوي والمكاني للدراسات المكشفة والمدرجة في قاعدة البيانات WoS في موضوع إدارة البيانات البحثية؟
5- ما أكثر المؤلفين إنتاجًا لأعمال مكشفة في قاعدة البيانات WoS في موضوع إدارة البيانات البحثية؟
6- ما أكثر الأعمال المستشهد بها من بين الأعمال المكشفة والمدجة في قاعدة البيانات WoS في موضوع إدارة البيانات البحثية؟
7- ما التخصصات ذات العلاقة بموضوع إدارة البيانات البحثية؟

3/1) أهداف الدراسة:

تهدف الدراسة بشكل عام إلى تحليل الأعمال المكشفة في قاعدة بيانات ISI Web of Science في موضوع إدارة البيانات البحثية، و رسم خريطة لشبكة التخصصات ذات العلاقة ببحوث إدارة البيانات البحثية، أما الأهداف التفصيلية فتتمثل في:

1. تحديد نوع الوثائق المكشفة في قاعدة البيانات في موضوع إدارة البيانات البحثية.
2. معرفة التوزيع المكاني للوثائق المكشفة في قاعدة البيانات في موضوع إدارة البيانات البحثية.
3. معرفة أكثر الأعوام التي شهدت نشر أعمال في مجال إدارة البيانات البحثية طبقا لمخرجات قاعدة البيانات.
4. التوزيع اللغوي للوثائق المكشفة في قاعدة البيانات في موضوع إدارة البيانات البحثية.
5. تحديد أكثر المؤلفين إنتاجًا لأعمال مكشفة في قاعدة البيانات WoS في موضوع إدارة البيانات البحثية. وكذلك الأعمال التي حصلت على نسبة عالية من الاستشهاد بها.
6. تحديد الموضوعات المرتبطة بموضوع إدارة البيانات البحثية من خلال فحص الكلمات المفتاحية الأكثر ترددًا في الأعمال المدرجة في قاعدة البيانات.

4/1) منهج الدراسة:

تم الاعتماد على أسلوب القياسات الببليومترية، وهي قياسات كمية وإحصائية توضح اتجاهات النشر وأنماط المساهمة في مجال معين من المجالات. والبيانات التي تم الحصول عليها من خلال التحليل الببليومتري لها قيمة كبيرة وتساعد الباحثين على إجراء مقارنات بين مختلف الباحثين والمؤسسات والدول واللغات.. الخ. كما أنها تلعب دورًا مهمًا في التعرف على نقاط الضعف والقوة في أي مجال من المجالات وتساعد صانعي السياسات على تحديد اتجاهاتهم المستقبلية.

5/1) محددات الدراسة:

تم البحث في قاعدة البيانات ISI Web of Science يوم 2 أبريل 2022 باستخدام مصطلح "Research Data Management" في حقل الكلمات المفتاحية، ويغطي الإطار الزمني كل السنوات حتى 2022م، وبكل أشكال الوثائق المتاحة ولم يتم الاقتصار على شكل معين،

ومن ثم بلغت نتائج البحث المسترجعة 594 وثيقة، قامت الباحثة بتصدير بياناتها إلى ملف إكسيل لتتمكن من فحص البيانات وتحليلها. كما تم الاستعانة ببرنامج VOSViewer لإجراء التصور الخاص بالشبكة متعددة التخصصات في أبحاث إدارة البيانات البحثية.

6/1) الدراسات السابقة:

تعد دراسة كل من (Jiming Hu وYin Zhang) عام 2017م من الدراسات القريبة من الدراسة الحالية التي اعتمدت على الأساليب الببليومترية وأدوات تحليل وتصور مختلفة للشبكات الاجتماعية لفحص هيكل وأنماط التخصصات العلمية المتداخلة مع موضوع البيانات الضخمة؛ حيث استخدمت الدراسة الفئات الموضوعيةsubject categories للمنشورات في قاعدة بيانات WoS كأساس لبناء شبكات التواجد المشترك من التخصصات واستخدمت الدراسة برنامج VOSviewer لعمل تصور أو رسم بياني شاملا للمجتمع متعدد التخصصات. وأشارت النتائج إلى اتساع نطاق التخصصات المشاركة في أبحاث البيانات الضخمة، ويميل التعاون العام بين التخصصات إلى التركيز في العديد من المجالات الرئيسية. فوفقًا لمؤشرات الشبكة، تعد علوم الكمبيوتر والهندسة والأعمال والاقتصاد من أهم المساهمين في أبحاث البيانات الضخمة، نظرًا لموقعهم ودورهم في شبكة التعاون البحثي[6].

أجرى الباحثان (Xu & Yu) في عام 2019م دراسة ببليومترية لتحليل منشورات البيانات الضخمة المكشفة في كشاف استشهادات العلوم Science Citation Index (SCI)، وكشاف استشهادات العلوم الاجتماعية Social Science Citation Index (SSCI). كان أحد التحليلات التي تم إجراؤها هو تحليل منشورات البيانات الضخمة وفقًا للبلدان والدوريات والمؤسسات البحثية الأكثر مساهمة. ومن ناحية أخرى، تم فحص ظهور الكلمات الرئيسية ومؤلفي المنشورات ونقاط البحث الساخنة الحالية واتجاهات التطوير المستقبلي. وأظهرت النتائج أن هناك 10989 وثيقة بين عامي 2009 و2018، بمتوسط 1099 مقالة منشورة كل عام. كان عدد المراجع والمؤلفين لكل ورقة يتزايد عامًا بعد عام. وبلغت نسبة التعاون الدولي في عام 2012 حدًا أقصى قدره 44.44٪، وكان معدل النمو الدولي في السنوات الأخرى أقل من 40٪. وجاءت جامعة كاليفورنيا أكثر المؤسسات إنتاجية، يليها الأكاديمية الصينية للعلوم. أما الدولة الأسرع نموًا كانت إيطاليا وتليها كوريا الجنوبية. وكانت دورية IEEE Access هي الأكثر

إنتاجية في مجال البيانات الضخمة. كما تم في هذه الدراسة، فحص ظهور الكلمات الرئيسية ومؤلفي المنشورات والموضوعات الحالية الساخنة واتجاهات التنمية المستقبلية[7].

وقدم (Tupan وKamaludin) في عام 2020 دراسة حول الأعمال المنشورة في مجلات الوصول الحر في مجال إدارة البيانات البحثية، وقد اعتمد على قاعدة بيانات Scopus لمعرفة عدد الأعمال المكشفة بها في هذه المجال، وقام بتحليلها حسب سنة النشر، ومصدر النشر، والمؤلفين، والمؤسسات، والبلدان ، وأنواع الوثائق ووكالات التمويل. كما رسم خرائط لإدارة البحث بناءً على الكلمات الرئيسية. أظهرت نتائج الدراسة أن عدد مصادر الوصول الحر لمنشورات إدارة البيانات البحثية قد بدأ منذ عام 1981 حتى وصل في عام 2019 إلى 49 منشورًا. وأن أكثر المجلات انتشارًا التي تتيح الوصول إلى إدارة بيانات البحث هي مجلة Data Science Journal، وأن المؤلفين الأكثر إنتاجية في أعمال إدارة البيانات البحثية هما Cox, A.M. and Pinfield, S، وبالنسبة للمؤسسات كانت أكبر المؤسسات التي ساهمت في نشر إدارة بيانات البحوث ذات الوصول المفتوح هي جامعة تورنتو وجامعة نيويورك. وكانت الدول الأكثر مساهمة هي الولايات المتحدة بـ 50 مطبوعة ، ثم الصين بـ 38 مطبوعة. المؤسسات التي قدمت أكبر تمويل مؤسستان هما: مؤسسة الأبحاث الألمانية Deutsche Forschungsgemeinschaft والمؤسسة القومية للعلوم the National Science Foundation[8].

وفي عام 2021م قدم (PICCOLO, D. M) وزملاؤه دراسة هدفت إلى تحديد الإنتاج العلمي الحالي حول موضوع "جودة البيانات في إدارة البيانات البحثية" من خلال قياس المؤشرات الببليومترية. وقد اعتمدت الدراسة على قواعد بيانات Web of Science و Scopus لتكوين مجموعة مواد الدراسة الببليومترية. وتم تحديد مجموعة من 77 مقالًا في الفترة بين عامي 1984 و 2020 ، حيث كان عام 2019 هو العام الذي يحتوي على العدد الأكبر من الأعمال المنشورة. وكانت الولايات المتحدة هي الدولة التي نشرت العدد الأكبر من الأوراق، بإجمالي 34 مقالة. وجاء مجال علوم الكمبيوتر أكثر إنتاجية حول هذا الموضوع ويشكل اتجاهًا في تعدد التخصصات مع العلوم البيولوجية والاجتماعية والصحية التطبيقية[9].

هذه الدراسات السابقة هدفت إلى دراسة الإنتاج الفكري المنشور في موضوع إدارة البيانات البحثية أو تحليل شبكات الكلمات المفتاحية لرصد التخصصات ذات الصلة في هذا الموضوع، ولكن هناك نوع آخر من الدراسات التي تتناول تحليل سجلات البيانات البحثية

نفسها التي يعتمد عليها المؤلفون، وليس مخرجات البحوث. ومن هذه الدراسات: قدم (Onyancha) في عام 2016م دراسة حاول فيها استكشاف حالة مشاركة البيانات البحثية بين الباحثين في أفريقيا جنوب الصحراء الكبرى (sub-Saharan Africa(SSA وعلى الصعيد الدولي. وتم استخراج البيانات ذات الصلة من كشاف الاستشهادات من البيانات (DCI) Data Citation Index باستخدام استراتيجية بحث متقدمة، حددت سنوات النشر بين عامي 2009 و 2014. وتم تحليل سجلات البيانات حسب الدولة، والمؤسسة، وفئة الموضوع، وسنة النشر. ونوع الوثيقة بالإضافة إلى عدد الاستشهادات. تم استخدام تحليل ارتباط سبيرمان لقياس العلاقة بين سجلات البيانات والمقالات البحثية. أشارت النتائج إلى أن 20 دولة فقط (من أصل 50) في إفريقيا جنوب الصحراء أنتجت سجل بيانات واحدًا على الأقل في مؤشر البيانات العالمي (DCI) ، وجاءت جنوب إفريقيا في صدارة مجموعة الدول بعدد 539 سجلاً (61.39٪) تليها كينيا والكاميرون وغانا. كما أظهرت النتائج أيضًا أن SSA تسهم بنسبة 0.03٪ فقط من البيانات البحثية مقارنة بـ 1.4٪ من المقالات البحثية العالمية. تعد المؤسسات البحثية والجامعات المساهمين الرئيسيين في بيانات البحث، والتي تركز على علم الوراثة (61.3٪) والكيمياء الحيوية والبيولوجيا الجزيئية (61.3٪) والزراعة (29.2٪) والغابات (27.3٪). أما عن الاستشهادات المرجعية، اجتذبت بيانات البحث متوسط استشهادات أقل من المقالات. يكشف التحليل الترابطي للبيانات عن وجود علاقة ارتباط ذات دلالة إحصائية بين نشر البيانات الأولية والمقالات البحثية نفسها[10].

وفي عام 2020م قدم (محمود شريف زكريا) بدراسة حاول فيها تحليل البيانات البحثية الأولية الموجودة فيعينة من الدوريات المصرية في مجالي العلوم البحتة والتطبيقية المكشفة في قاعدة البيانات (ISI Web of Science)، والمنشورة عام 2018م، وبلغ عدد الدوريات 31 دورية، بينما عدد المقالات المصحوبة بمواد تكميلية 622 مقالة، تضمنت 866 ملف تكميلي للبيانات البحثية الأولية. وكان من أهم نتائج الدراسة أن البيانات البحثية توزعت على خمس فئات هي: الإيضاحيات، والاحصائيات، والبيانات المختلطة، والبيانات النصية، والوسائط المتعددة. أما صيغ هذه الملفات فتوزعت على صيغة MS Word، وصيغة PDF. وأنها ارتبطت بأقسام مختلفة في المقالات هي أقسام النتائج، والمنهج، والمناقشة[11].

ومن الواضح أن هاتين الدراستين لم تتناولا إدارة البيانات البحثية، وإنما ركزتا على تحليل ملفات البيانات البحثية، فالدراسة الأولى قامت بتحليل مجموعات البيانات البحثية نفسها المكشفة في قاعدة البيانات باستخدام أداة تتيجها قاعدة البيانات تقوم بتحليل مجموعات البيانات البحثية. والدراسة الثانية قامت بتحليل المواد التكميلية الموجودة مع المقالات المنشورة في الدوريات المصرية عام 2018 في مجالي العلوم البحتة والتطبيقية.

7/1) التعريف بقاعدة بيانات ISI WoS:

تعد قاعدة بيانات شبكة العلوم Web of Science أقدم قاعدة بيانات للاستشهادات المرجعية إذ تضم استشهادات مرجعية وبيانات ببليوجرافية تعود للعام 1900م، وقد أنتجها معهد تومسون رويترز للمعلومات العلمية (ISI) اعتمادًا على كشاف استشهادات العلوم Science Citation Index الذي أنشأه يوجين جارفيلد في ستينيات القرن العشرين.[12] وتتكون من سبع قواعد بيانات استشهادات مرجعية، بما في ذلك المعلومات المختلفة التي تم جمعها من الدوريات والمؤتمرات والتقارير والكتب وسلاسل الكتب. أما القواعد السبع فهي: كشاف استشهادات العلوم الاجتماعية Social Sciences Citation Index (SSCI)، وكشاف استشهادات العلوم الموسع Science Citation Index Expanded (SCI Expanded)، وكشاف استشهادات العلوم الإنسانية والآداب Arts and Humanities Citation Index (A&HCI)، وكشاف استشهادات الكتب Book Citation Index (BCI)، وكشاف استشهادات أعمال المؤتمرات Conference Proceedings Citation Index (CPCI)، وكشاف الكيمياء Index Chemicus (IC)، وقاعدة أخرى للتفاعلات الكيميائية تسمى Current Chemical Reactions (CCR)[13].

كانت تدار سابقا من قبل الناشر تومسون رويترز منذ بداية تسعينات القرن العشرين، .ويديرها حاليا Clarivate Analytics منذ عام 2016م التي تصدر معامل تأثير الدوريات العلمية. وتحتوي مجموعة Web of Science Core Collection على أكثر من 21100 مجلة علمية ذات جودة عالية، تم نشرها في جميع أنحاء العالم (بما في ذلك المجلات ذات الوصول الحر) في أكثر من 250 تخصصًا من العلوم والعلوم الاجتماعية والفنون والعلوم الإنسانية. بالإضافة غلى أعمال المؤتمرات والكتب[14].

ومن الجدير بالذكر أن منصة شبكة العلوم WoS تستضيف الكشاف العربي للاستشهادات المرجعية، الذي يساهم في جعل المجلات العربية في متناول الباحثين في جميع أنحاء العالم من خلال ربط المحتوى العلمي العربي بشبكة الاستشهادات العالمية من المجلات العلمية عالية الجودة المُحكّمة مع أكثر من 1.7 مليار مرجع من المراجع المستشهد بها. وقد تم بناء الكشّاف العربي للاستشهادات المرجعية من قبل كلاريفيت Clarivate وبتمويل من الحكومة المصرية، ويمكن البحث في الكشّاف العربي للاستشهادات المرجعية باللغتين العربية والإنجليزية[15].

2) الإطار النظري للدراسة:

1/2) البيانات البحثية:

هي البيانات التي تم إنشاؤها أثناء أنشطة البحث، وهي موارد قيمة يمكن الوصول إليها وتصفحها واستشارتها واستخدامها والبناء عليها للأغراض الأكاديمية والبحثية والعلمية، وذلك إذا تم تخزينها بشكل صحيح. ويمكن للباحثين إضافة وجهات نظرهم التي قد تساعد في توليد معرفة جديدة. وبالتالي يمكنهم تجنب إهدار الوقت والجهود في جمع البيانات من جديد[16].

وتتضمن البيانات البحثية كل جزء من البيانات تم الحصول عليه أثناء عملية البحث، ويمكن أن تتخذ أشكالا متعددة منها: النصوص، وجداول البيانات، والاستبانات، والصور، والأفلام، والشرائح، والملاحظات المعملية، والإحصاءات، والملاحظات، ونتائج التجارب، والقياسات، والعينات، والخوارزميات... الخ[17].

وللبيانات البحثية أهمية كبيرة في الاتصال العلمي، فهي تفيد الباحثين المبتدئين في التعرف على الأساليب العلمية المختلفة التي اتبعها المؤلفون السابقون في معالجة الظواهر العلمية، من قبيل كيفية إعداد أدوات جمع البيانات، أو طرق تصميم التجارب المعملية، أو الأساليب الإحصائية المستخدمة في تحليل البيانات. ومن جانب آخر فهي تزيد من احتمالية رفع معدل الاستشهاد بالمقالات التي تتيح البيانات البحثية الأولية، كما أنها تساعد المحكمين والقراء المتخصصين على التأكد من مصداقية نتائج الأبحاث العلمية، وتعزيز فرص قبول نشر المقالات من جانب هيئات تحرير الدوريات العلمية[18]. وكذلك تعميق العلاقات بين الباحثين

والتعاون فيما بينهم وتبادل البيانات البحثية مما يسرع من وتيرة الإنتاج العلمي في المجالات المختلفة.

ويضاف إلى ذلك أن بيانات البحث يمكن أن تكون أساسًا لأسئلة بحثية أخرى وجديدة، حيث يمكن أن تكون جزءًا لا يتجزأ من دورة حياة البحث منذ البداية، ويمكن إعادة تحليل بيانات البحث باستخدام طرق رقمية جديدة ومبتكرة لم تكن معروفة وقت الحصول على البيانات. وكذلك يمكن استخدامها من قِبل التخصصات الأخرى، وبالتالي تشجيع البحث متعدد التخصصات[19].

2/2) إدارة البيانات البحثية:

معظم البيانات البحثية غالبًا ما يكون لها عمر أطول بكثير من المشروع البحثي الذي تم إنجازه اعتمادًا عليها، وقد يستمر الباحثون في العمل مع البيانات وتحليلها وإعادة استخدامها، كما قد يتم ذلك من قِبل باحثين آخرين. لذا فقد بدأ الحديث منذ العقد الأخير من القرن العشرين عن إدارة البيانات البحثية لتعزيز دورة حياة البيانات البحثية كمفهوم يدعم ممارسات الحفاظ على البيانات وتنظيمها، وتوثيقها وتمكين استدامتها ومشاركتها[20].

تعرف إدارة البيانات البحثية بأنها طريقة تمكن من تكامل وتنظيم وتشغيل البيانات التي تم إنشاؤها أثناء إجراء البحث العلمي، مثل إنتاج البيانات والوصول إليها والتحقق منها وتخزينها وإعادة استخدامها بمساعدة مجموعة من الأدوات المساعدة سهلة الاستخدام في البيئة الافتراضية[21]. وفي تعريف مشابه للباحثين (Angus & Tedds) إدارة البيانات البحثية هي نشاط يتعلق بكيفية إدارة الأفراد للبيانات أثناء الأنشطة البحثية أو الكتابة العلمية، وتشمل هذه الأنشطة: الجمع والتنظيم والإدارة والتخزين والأمن، والنسخ الاحتياطي والصيانة وتبادل البيانات[22].

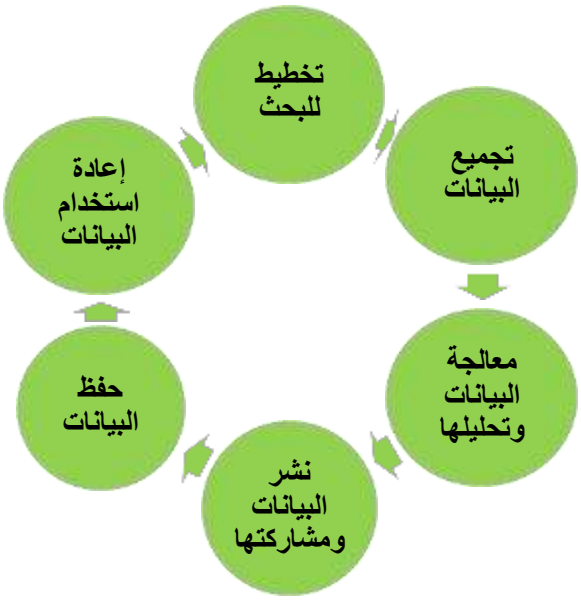
وتهدف إدارة البيانات البحثية إلى جمع والتقاط وتخزين وتتبع وأرشفة جميع البيانات التي يتم إنتاجها في المشاريع والتجارب العلمية. إلى جانب هذه البيانات، يجب تخزين جميع خطوات المعالجة على هذه البيانات؛ مما يؤدي في النهاية إلى إنتاج منشورات علمية جديدة[23].

وبالتالي فإن إدارة البيانات البحثية موجودة في جميع مراحل البحث؛ نظرًا لأنها تتضمن عمليات إنشاء البيانات، والتخطيط لإعادة الاستخدام، والمعالجة، وتنظيم البيانات

نفسها في هيكلها وتحليل محتواها، وحفظها، كل هذا مع مراعاة الأمان في عمليات الوصول إليها وتخزينها واسترجاعها ومشاركتها وإعادة استخدامها.

3/2) دورة حياة البيانات البحثية:

وضعت منظمة Data Archive التي أنشئت في عام 1967م، والتابعة لجامعة Essex في المملكة المتحدة، تصورًا لدورة حياة البيانات البحثية يوضحها الشكل التالي:



شكل رقم (2) دورة حياة البيانات البحثية

وفقًا للشكل السابق، فإن دورة حياة البيانات البحثية تشتمل على المراحل التالية[24]:

1- التخطيط للبحث: وتتضمن هذه المرحلة: تصميم البحث، والتخطيط لإدارة البيانات، والتخطيط لتجميع البيانات ومعالجتها، والكشف عن مصادر البيانات الحالية، والتخطيط لمشاركة البيانات.

2- تجميع البيانات وتتضمن هذه المرحلة: جمع البيانات، وتمثيل البيانات البحثية بالبيانات الوصفية.

3- معالجة البيانات وتحليلها: وتتضمن هذه المرحلة: إدخال البيانات ورقمنتها ونسخها وترجمتها، وفحصها، واستخلاصها، وصفها وتوثيقها، وتخزينها، وتحليلها، وكيفية الاستشهاد بها.

4- نشر البيانات ومشاركتها وتتضمن هذه المرحلة: إنشاء وثائق للمستفيد، وإنشاء البيانات الوصفية المكشفة، واختيار الدخول الملائم للبيانات، ومشاركة البيانات، وتسويق البيانات.

5- حفظ البيانات وتتضمن هذه المرحلة: وضع البيانات في أفضل تنسيق، وعمل نسخ احتياطية للبيانات وتخزينها، بغرض حفظها.

6- إعادة استخدام البيانات وتتضمن هذه المرحلة: إجراء تحليل ثانوي للبيانات، والقيام بمتابعة المشروع البحثي، وإجراء مراجعات للمشروع البحثي، والتأكد من دقة النتائج، واستخدام البيانات في التدريس والتعلم.

3) نتائج الدراسة التحليلية:

1/3) التوزيع الزمني للأعمال المدرجة في WoS في موضوع إدارة البيانات البحثية:

رصدت قاعدة البيانات WoS أول دراسة عن إدارة البيانات البحثية عام 1974، وبالتالي يعود بداية الحديث عن إدارة البيانات البحثية إلى ثمانية وأربعين عامًا مضت. ويوضح الجدول التالي التوزيع الزمني للدراسات الخاصة بإدارة البيانات البحثية على مدار هذه السنوات.

جدول رقم (1) التوزيع الزمني للأعمال في موضوع إدارة البيانات البحثية

السنة	العدد	النسبة المئوية	السنة	العدد	النسبة المئوية
2022	23	3.87	2008	1	0.17
2021	77	12.96	2007	3	0.51
2020	79	13.30	2004	1	0.17
2019	88	14.81	2003	1	0.17
2018	73	12.29	2001	1	0.17
2017	74	12.46	2000	1	0.17
2016	42	7.07	1997	1	0.17

السنة	العدد	النسبة المئوية	السنة	العدد	النسبة المئوية
2015	45	7.58	1993	1	0.17
2014	43	7.24	1984	1	0.17
2013	14	2.36	1983	1	0.17
2012	17	2.86	1982	1	0.17
2010	1	0.17	1981	1	0.17
2009	3	0.51	1974	1	0.17
المجموع				**594**	**100.00**

يلاحظ من الجدول السابق أن أعلى نسبة للنشر في موضوع إدارة البيانات البحثية كانت في عام 2019م إذ بلغت الدراسات المنشورة (88) عملا وهي أعلى نسبة على الإطلاق. كما يلاحظ أن أول عمل في هذه الموضوع ظهرت عام 1974م وكان بعنوان "نظام تفاعلي قابل للتكيف لإدارة البيانات الطبية والبحثية ADAPTABLE INTERACTIVE SYSTEM FOR MEDICAL AND RESEARCH DATA MANAGEMENT" نشرت في دورية METHODS OF INFORMATION IN MEDICINE ، واستمر نشر عمل واحد حتى عام 2010 ــ باستثناء عامي 2007م، و2009م فقد نشر ثلاثة أعمال في كل منهما ـــ ثم حدثت طفرة كبيرة في عام 2012م حيث نشر سبعة عشر عملا، ثم استمرت الزيادة عامًا بعد عام.

2/3) التوزيع تبعًا لنوع مادة النشر:

تنوعت الأعمال في موضوع إدارة البيانات البحثية ما بين مقالات دوريات، وأعمال مؤتمرات، ومراجعات علمية، وفصول كتب، وعروض كتب، ومواد تحريرية، وملخصات اجتماعات، ويوضح الجدول التالي توزيع العمال المدرجة في قاعدة بيانات WOS على هذه الأنواع:

جدول قم (2) توزيع الأعمال المدرجة في قاعدة البيانات على نوع المواد

نوع المادة	العدد	النسبة المئوية
مقالات دوريات	379	63.80
أعمال مؤتمرات	118	19.87
مراجعات علمية	30	5.05
المواد ذات الوصول المبكر (*)	21	3.54

نوع المادة	العدد	النسبة المئوية
مواد تحريرية	10	1.68
عروض كتب	12	2.02
فصول كتب	11	1.85
ملخصات اجتماعات	12	2.02
أوراق بيانات	1	0.17
المجموع	**594**	**100**

. يلاحظ من الجدول السابق أن المقالات المنشورة في الدوريات هي أكثر أشكال مصادر المعلومات المدرجة في قاعدة البيانات WOS في موضوع إدارة البيانات البحثية. وبلغ عددها 379 مقالة بنسبة 63.8% من إجمالي عدد المواد.

ولوحظ أيضًا أن قاعد البيانات WOS قد أتاحت واحد وعشرين عملا تحت مسمى "المواد ذات الوصول المبكر Early Access" وهي مواد مفهرسة بالكامل في قاعدة بيانات WOS باستثناء تاريخ النشر النهائي والتعيين النهائي للمجلد والإصدار ورقم الصفحة فهي مقبولة للنشر ولكنها لم تنشر بشكل نهائي.

كما لوحظ أن أحد الأعمال عبارة عن بيانات بحثية بالفعل وهي ورقة في علم البلازما بعنوان Plasma-MDS, a metadata schema for plasma science with examples from plasma technology نشرت عام 2020م في مجلة Scientific Data، وهي عبارة عن مخطط للبيانات الوصفية metadata schema ويستكمل بالبيانات الرقمية التي تم الحصول عليها من خلال إجراءات التشخيص المطبقة من قبل الباحثين.

أما عروض الكتب فقد بلغت 12 عرضًا ، منهم ستة عروض لكتاب واحد هو كتاب لمؤلفه بيتر لوند Peter Lund نشره عام 2015 بعنوان: "تقديم خدمات إدارة البيانات البحثية: أساسيات الممارسة الجيدة :Delivering Research Data Management Services Fundamentals of Good Practice" وهو دليل على أهمية ذلك الكتاب.

3/3) التوزيع حسب مجال البحث:

بلغ عدد المجالات البحثية التي نشر بها أعمال في إدارة البيانات البحثية خمسة وسبعين مجالا بحثيًا، مما يعني أن البيانات البحثية في جميع المجالات العلمية والبحثية يمكن إدارتها، ويوضح الجدول التالي توزيع الأعمال المنشورة على هذه المجالات:

جدول قم (3) توزيع الأعمال المدرجة في قاعدة البيانات على المجال البحثي

المجال البحثي	العدد	النسبة المئوية
علم المكتبات والمعلومات	310	52.19
علم الحاسب	221	37.21
المعلوماتية الطبية	33	5.56
خدمات علوم الرعاية الصحية	26	4.38
تكنولوجيا العلوم	24	4.04
الهندسة	22	3.70
الكيمياء	19	3.20
الكيمياء الحيوية والبيولوجيا الجزيئية	16	2.69
التعليم	16	2.69
البيولوجيا الرياضية	12	2.02
الجغرافيا الطبيعية	9	1.52
العلوم الاجتماعية	9	1.52
الإنسانيات	8	1.35
الاستشعار من بعد	8	1.35
اقتصاديات الأعمال	7	1.18
العلوم البيئية	6	1.01
علوم الحياة	6	1.01
عدد (5) مجالات	5 لكل منهم	0.84
علم الصيدلة	4	0.67

المجال البحثي	العدد	النسبة المئوية
عدد (8) مجالات	3 لكل منهم	0.51
عدد (13) مجال	2 لكل منهم	0.34
عدد (31) مجالا	1 لكل منهم	0.17
المجموع	**837**	

وصل العدد الإجمالي (837) لأن بعض الأعمال صنفت في أكثر من مجال بحثي واحد.

ويلاحظ من الجدول السابق أن مجال المكتبات والمعلومات من أكثر المجالات الموضوعية التي نشر فيها أعمال في موضوع إدارة البيانات البحثية، حيث بلغ عدد الأعمال المدرجة في قاعدة البيانات WOS (310) عملا في مجال المكتبات والمعلومات بنسبة 52.19% من عدد الأعمال، يليه علم الحاسب بنسبة 37.21%.

4/3) التوزيع المكاني للأعمال المدرجة في قاعدة البيانات:

تتوزع الأعمال المدرجة في قاعدة البيانات WoS في مجال إدارة البيانات البحثية على 73 دولة من دول العالم، كما يتضح من الجدول التالي:

جدول رقم (4) التوزيع المكاني للأعمال المكشفة في WoS

الدولة	العدد	% من عدد الأعمال	الدولة	العدد	% من عدد الأعمال
الولايات المتحدة الأمريكية	138	23.23	السويد	7	1.18
ألمانيا	125	21.04	فنلندا	5	0.84
إنجلترا	65	10.94	نيوزيلندا	5	0.84
هولندا	34	5.72	أيرلندا	4	0.67
كندا	32	5.39	النرويج	4	0.67
استراليا	25	4.21	كوريا الجنوبية	4	0.67
جنوب افريقيا	20	3.37	بلغاريا	3	0.51
سويسرا	20	3.37	استونيا	3	0.51
البرتغال	18	3.03	اليونان	3	0.51
الهند	17	2.86	كينيا	3	0.51

الدولة	العدد	% من عدد الأعمال	الدولة	العدد	% من عدد الأعمال
الصين الشعبية	17	2.86	مالاوي	3	0.51
البرازيل	16	2.69	نيجيريا	3	0.51
فرنسا	14	2.36	بولندا	3	0.51
ايطاليا	13	2.19	روسيا	3	0.51
أسبانيا	11	1.85	السعودية	3	0.51
النمسا	10	1.68	جمهورية التشيك	2	0.34
اسكوتلندا	10	1.68	الدنمارك	2	0.34
المجر	8	1.35	غانا	2	0.34
اليابان	8	1.35	اندونيسيا	2	0.34
باكستان	8	1.35	سنغافورة	2	0.34
تركيا	8	1.35	سلوفاكيا	2	0.34
بلجيكا	7	1.18	أوكرانيا	2	0.34
29 دولة هي: (كرواتيا – كوبا- قبرص – مصر – ايسلندا – إيران – العراق – اسرائيل – الأردن – كازاخستان – الكويت – لبنان – ليتوانيا – ماليزيا – لوكسمبورج – مالطة – المكسيك – ميانمار – الفلبين – الكونغو – رومانيا – السنغال – سلوفينيا – تايوان – ترينداد وتوباجو – اوغندا – فييتنام – ويلز – زيمبابوي)				1 لكل دولة	0.17
المجموع*				**723**	

* وصل العدد إلى (723) لوجود تأليف مشترك لباحثين من دول مختلفة (فرق بحثية)

يتضح من الجدول السابق تفوق الولايات المتحدة الأمريكية في عدد الأعمال المدرجة في قاعدة WoS في مجال إدارة البيانات البحثية، يليها مباشرة ألمانيا ثم انجلترا ثم هولندا، وعدد الأعمال التي تنتمي لهذه الدول الأربع يمثل 50 % من إجمالي الأعمال، أو بصورة أخرى يساوي إجمالي الأعمال المنتمية لبقية الدول البالغ عددها 69 دولة.

أن 29 دولة لديهم عمل واحد فقط مدرج في قاعدة WoS في مجال إدارة البيانات البحثية، وسبع دول لديهم عملان، وتسع دول لديهم ثلاثة أعمال.

بلغ عدد الأعمال التي قام بها باحثون ينتمون للدول العربية ثمانية أعمال موزعة على ست دول عربية، هي: السعودية (ثلاثة أعمال)، وكل من مصر والعراق والأردن والكويت ولبنان (عمل واحد لكل منهم)

5/3) التوزيع اللغوي للأعمال المدرجة في قاعدة البيانات:

تتوزع الأعمال المدرجة في قاعدة البيانات WoS في موضوع إدارة البيانات البحثية على عشر لغات يوضحها الجدول التالي:

جدول رقم (5) التوزيع اللغوي للأعمال المكشفة في WoS

م	اللغة	العدد	النسبة المئوية %
.1	الإنجليزية	532	89.56
.2	الألمانية	39	6.57
.3	البرتغالية	12	2.02
.4	الأسبانية	4	0.67
.5	التركية	2	0.34
.6	الفرنسية	1	0.17
.7	الإيطالية	1	0.17
.8	اليابانية	1	0.17
.9	الروسية	1	0.17
.10	الأوكرانية	1	0.17
المجموع		**594**	**100**

يتضح من الجدول السابق أن اللغة الإنجليزية تتصدر لغات النشر؛ حيث وجد أن 89.56 % من الأعمال المنشورة والمدرجة في قاعدة البيانات WoS باللغة الإنجليزية، بينما نشر 6.56% باللغة الألمانية، و2.02% باللغة البرتغالية. وبالطبع لا يوجد مواد منشورة باللغة العربية.

6/3) المؤلفون الأعلى إنتاجية:

يوضح الجدول التالي أسماء المؤلفين في المراتب الخمسة الأولى من حيث الإنتاجية العلمية في موضوع إدارة البيانات البحثية.

جدول رقم (6) المؤلفون الأعلى إنتاجية في موضوع إدارة البيانات البحثية.

الرتبة	المؤلف	H-Index(*)	العدد	النسبة المئوية
1	Da Silva JR	5	15	2.53
2	Ribeiro C	4	13	2.19
3	Cox AM	10	10	1.68
4	Castro JA	4	8	1.35
5	7 مؤلفين: (Amorim RC -Chard K -Foster I - Pinfield S - Koltay T - Read KB Surkis A)		6 لكل	1.01

(*) تم معرفة معامل هيرش H-Index للمؤلفين من خلال الآداة التي تتيحها قاعدة البيانات ISI WoS وعنوانها:

https://0810o74cv-1104-y-https-www-webofscience-com.mplbci.ekb.eg/wos/author/search

جاء المؤلف البرتغالي João Rocha da Silva. صاحب الـــ H-Index = 5 في الصدارة حيث ساهم بعدد خمسة عشر عملا، يليه في المرتبة الثانية البرتغالية Ribeiro, Cristina صاحبة الـــ H-Index =4 حيث ساهمت بعدد ثلاثة عشر عملا، وفي المرتبة الثالثة المؤلف البريطاني Andrew M. Cox صاحب الـ H-Index = 10 حيث ساهم بعشرة أعمال. ثم في المرتبة الرابعة المؤلف البرتغالي João Aguiar Castro صاحب الـــ H-Index = 4 حيث ساهم بثمانية أعمال. ثم جاء سبعة مؤلفين في المرتبة الخامسة حيث نشر كل مؤلف منفردًا ستة أعمال.

7/3)إنتاجية الدوريات المكشفة في ISI WoS في مجال إدارة البيانات البحثية:

يوضح الجدول التالي إنتاجية الدوريات العشرة الأولى المكشفة في ISI WoS في مجال إدارة البيانات البحثية.

جدول رقم (7) أعلى الدوريات الناشرة في مجال إدارة البيانات البحثية

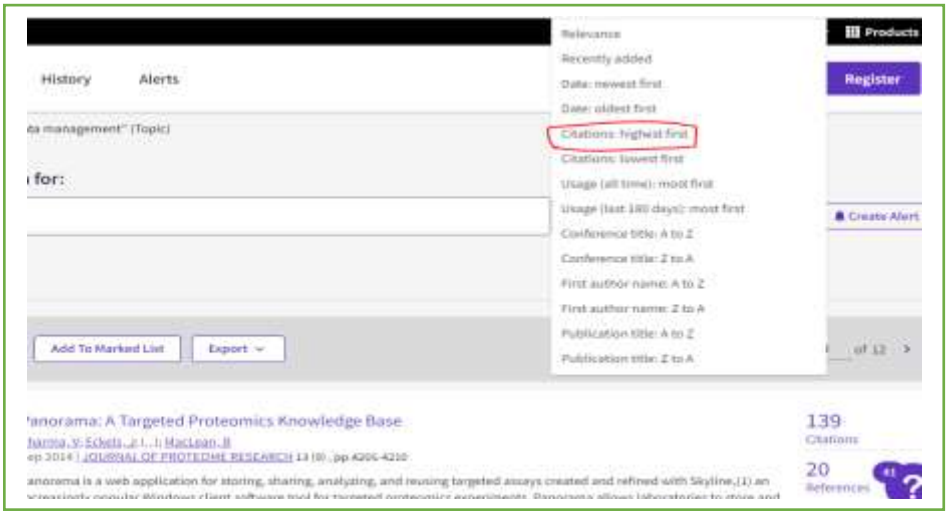
الرتبة	الدورية	العدد	النسبة
1	LECTURE NOTES IN COMPUTER SCIENCE	25	4.21
2	BIBLIOTHEK FORSCHUNG UND PRAXIS	18	3.03
2	COMMUNICATIONS IN COMPUTER AND INFORMATION SCIENCE	18	3.03
3	IFLA JOURNAL	16	2.69
4	PROCEDIA COMPUTER SCIENCE	13	2.19
5	JOURNAL OF ACADEMIC LIBRARIANSHIP	11	1.85
6	STUDIES IN HEALTH TECHNOLOGY AND INFORMATICS	10	1.68
7	INFORMATION WISSENSCHAFT UND PRAXIS	9	1.52
7	LIBRARY MANAGEMENT	9	1.52
8	DESIDOC JOURNAL OF LIBRARY INFORMATION TECHNOLOGY	8	1.35

جاءت دورية LECTURE NOTES IN COMPUTER SCIENCE في المرتبة الأولى ونشرت 25 عملا، وهي دورية متخصصة في مجال علم الكمبيوتر وتصدر باللغة الإنجليزية منذ عام 1973. وفي المرتبة الثانية دوريتان نشرت كل منهما 18 عملا هما: دورية BIBLIOTHEK FORSCHUNG UND PRAXIS باللغة الألمانية وتغطي جميع جوانب علم المكتبات، ودورية COMMUNICATIONS IN COMPUTER AND INFORMATION SCIENCE وهي سلسلة مخصصة لنشر وقائع مؤتمرات علوم الكمبيوتر . وفي المرتبة الثالثة دورية IFLA JOURNAL المتخصصة في علم المكتبات والمعلومات وهي دورية فصلية تصدر عن الاتحاد الدولي لجمعيات المكتبات ومؤسساتها، قامت بنشر 16 عملا. وفي المرتبة الرابعة دورية PROCEDIA COMPUTER SCIENCE التي تم إطلاقها في عام 2010م، وهي تنشر وقائع مؤتمرات عالية الجودة في مجال علوم الكمبيوتر وقامت بنشر 13 عملا، وفي المرتبة الخامسة دورية JOURNAL OF ACADEMIC LIBRARIANSHIP وهي دورية متخصصة في مجال المكتبات

الأكاديمية تصدر منذ عام 1975م ونشرت إحدى عشر عملا، وفي المرتبة السادسة دورية STUDIES IN HEALTH TECHNOLOGY AND INFORMATICS التي بدأت في عام 1990م بالتعاون مع برنامج الاتحاد الأوروبي، وتنشر البحوث في مجال المعلوماتية الصحية وقامت بنشر عشرة أعمال، وفي المرتبة السابعة دوريتان نشرت كل منهما تسعة أعمال هما: دورية INFORMATION WISSENSCHAFT UND PRAXIS وتصدر من 2012م باللغة الألمانية، ودورية LIBRARY MANAGEMENT وهي متخصصة في إدارة المكتبات تصدر منذ عام 1979م. وأخيرا دورية DESIDOC JOURNAL OF LIBRARY INFORMATION TECHNOLOGY وتصدر في الهند عن مركز التوثيق والمعلومات العلمية Defence Scientific Information & Documentation Centre (DESIDOC) منذ عام 1912.

8/3) المقالات الأكثر استشهادًا highly cited في مجال إدارة البيانات البحثية:

يوضح الجدول التالي بيانات المقالات العشرة الأولى الأكثر استشهادًا من قبل الباحثين، وتم تحديد ذلك من خلال قاعدة البيانات ISI WoS نفسها، إذ أنها تتيح خاصية ترتيب نتائج البحث في قاعدة البيانات وفق عدد من المعايير منها: الاستشهادات الأعلى، كما يظهر في الشكل التالي:



شكل رقم (3) معايير ترتيب نتائج البحث في قاعدة البيانات ISI WoS

وجاءت المقالات الأعلى استشهادًا كما يلي:

جدول رقم (8) المقالات الأعلى استشهادًا في مجال إدارة البيانات البحثية

الرتبة	مؤلف المقال	عنوان المقال	سنة النشر	عدد الاستشهادات
1	Sharma, V; Eckels, J; (...); MacLean, B	Panorama: A Targeted Proteomics Knowledge Base	2014	139
2	Corrall, S; Kennan, MA and Afzal, W	Bibliometrics and Research Data Management Services: Emerging Trends in Library Support for Research	2013	100
3	Tenopir, C; Sandusky, RJ; (...); Birch, B	Research data management services in academic research libraries and perceptions of librarians	2014	86
4	Anderson, NR; Lee, S; (...); Tarczy-Hornoch, P	Issues in biomedical research data management and analysis: Needs and barriers	2007	78
5	Cox, AM and Pinfield, S	Research data management and libraries: Current activities and future priorities	2014	76
6	Skripcak, T; Belka, C; (...); Baumann, M	Creating a data exchange strategy for radiotherapy research: Towards federated databases and anonymised public datasets	2014	64
7	Cox, AM; Kennan, MA; (...); Pinfield, S	Developments in Research Data Management in Academic Libraries: Towards an Understanding of Research Data Service Maturity	2017	62
8	Arend, D; Lange, M; (...); Scholz, U	e!DAL - a framework to store, share and publish research data	2014	52
9	Arend, D; Junker, A; (...); Lange, M	PGP repository: a plant phenomics and genomics data publication infrastructure	2016	44
10	Pinfield, S; Cox, AM and Smith, J	Research Data Management and Libraries: Relationships, Activities, Drivers and Influences	2014	44

يتضح من الجدول السابق أن ست مقالات من المقالات العشرة الأعلى استشهادًا في مجال إدارة البيانات البحثية صدرت في عام 2014م، بنسبة 60%، ومقالة واحدة في السنوات

2007م، و2013م، و2016م، و2017م. أما بالنسبة للمؤلفين فجاءت أعمال Cox, AM ضمن أكثر الأعمال استشهادًا بواقع 3 أعمال (30%)، ثم جاءت أعمال كل من Arend, D و ,Pinfield S وKennan, MA بواقع عملين (20%). كما يلاحظ أن جميع الأعمال العشرة ذات تأليف مشترك.

ومن حيث موضوعات المقالات الأعلى استشهادًا، جاءت المقالة الأولى التي تتناول تطبيق Panorama وهو تطبيق ويب لتخزين ومشاركة وتحليل وإعادة استخدام بيانات البحوث الطبية الحيوية، وجاء في المرتبة الأولى بواقع 139 استشهادًا، ويمكن تنزيل هذا التطبيق على الأجهزة الخاصة بالمستخدمين.

9/3) تحليل الكلمات المفتاحية:

أظهرت الدراسات أن استخدام الكلمات المفتاحية Keywords للوثيقة لتحديد الموضوع أكثر فاعلية من استخدام الكلمات والعبارات الواردة في العنوان أو المستخلص. لذا يحلل هذا القسم الكلمات المفتاحية لجميع المواد ذات الصلة، ويبني شبكات التواجد المشترك بينهاCo-occurrence of keywords . ومن أجل تحديد الاتجاهات بشكل أكبر، تم تقسيم الفترة من 2012م إلى 2022م وهي الفترة التي شهدت نموًا في الإنتاج الفكري في موضوع إدارة البيانات البحثية (راجع جدول رقم 1) إلى ثلاث مراحل مختلفة، وهي: 2012 - 2016، و2017 - 2019، و 2020 - 2022. هذا القسم يأخذ في الاعتبار بشكل رئيسي جانبين.

الأول: هناك المزيد من الدراسات في السنوات الأخيرة ودراسات أقل في السنوات السابقة، ومن المفيد موازنة كمية الدراسات في فترات زمنية مختلفة.

الثاني: يعد هذا التقسيم أكثر فائدة في تحديد النقاط الساخنة للبحوث الحديثة في هذا المجال في كل فترة من الفترات الثلاث.

ولأجل جعل الشبكات للفترات الزمنية المختلفة أكثر إيجازًا وسهولة في التعرف على المعلومات الأساسية، لا يتم عرض بعض الكلمات المفتاحية ذات التردد المنخفض في الشبكة. حيث تم تحديد أن تظهر الكلمات المفتاحية التي تتردد أربع مرات على الأقل في الشبكة.

وفي الشبكات المختلفة المعروضة فيما يلي يجب التأكيد على:

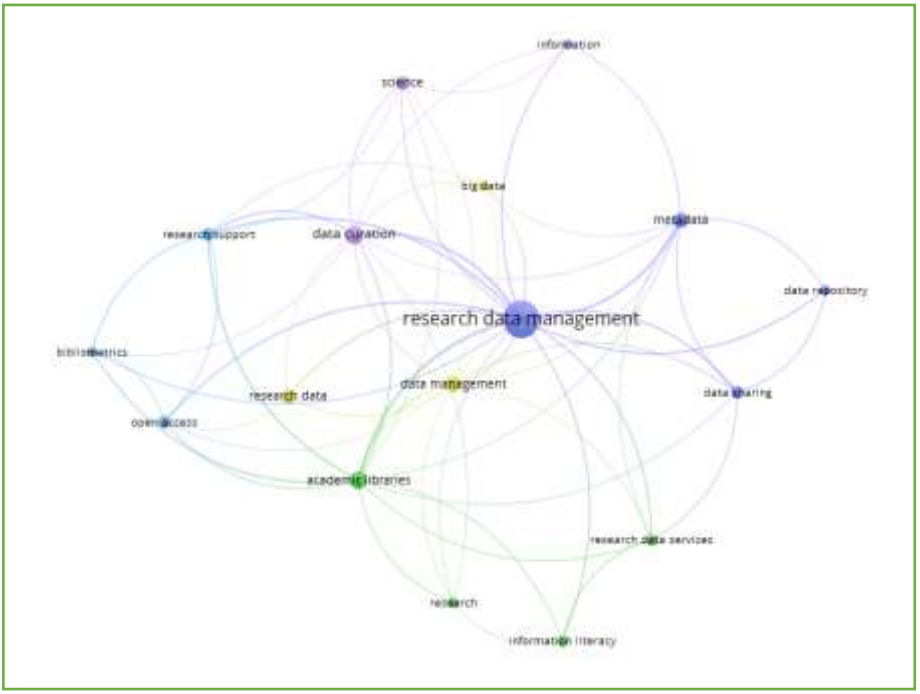
- تمثل العقدة الكلمة المفتاحية، وكلما كبرت العقدة، دل ذلك على زيادة عدد تكرارات الكلمة المفتاحية.

- ويمثل اتصال العقدتين كلمتين رئيسيتين تظهران معًا في الأعمال الخاصة بإدارة البيانات البحثية، ويمثل سمك الاتصال عدد التكرارات المشتركة.
- من الواضح أن أكبر عقدة من حيث الحجم في كل الشبكات تمثل مصطلح إدارة البيانات البحثية، حيث تميل كل الدراسات المسترجعة إلى استخدام هذا المصطلح في الكلمات المفتاحية بجانب مصطحات أخرى. وهو ما توضحه الأشكال التالية التي تمثل الفترات الزمنية المختلفة.

وفيما يلي عرض هذه الشبكات:

أولا :الفترة من 2012م إلى 2016م

يوضح الشكل التالي مدى تردد الكلمات المفتاحية الواردة في الدراسات المكشفة في قاعدة بيانات ISI WoS خلال الفترة 2012م – 2016م في موضوع إدارة البيانات البحثية.

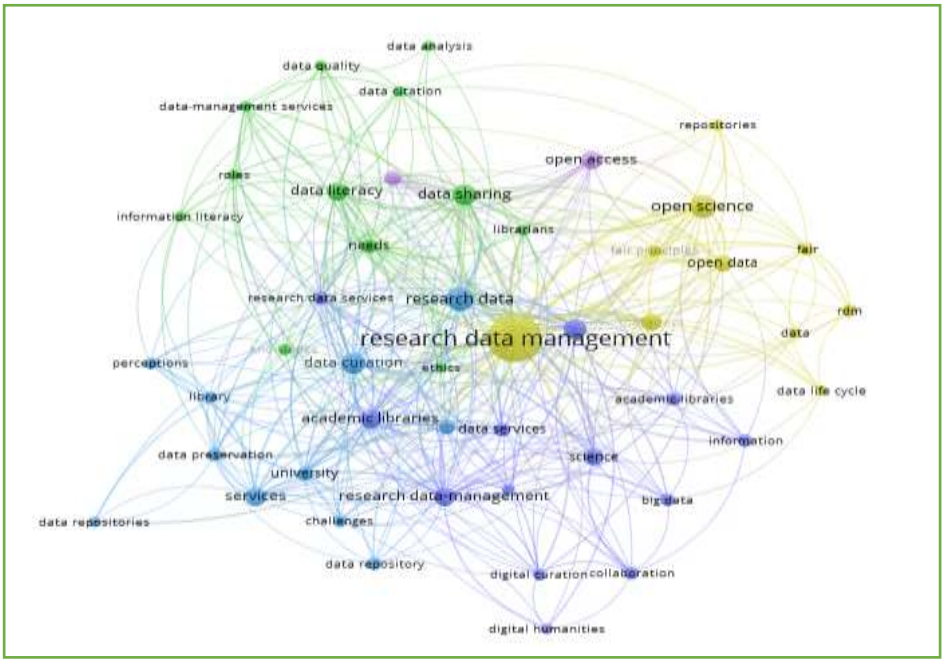


شكل رقم (4) تردد الكلمات المفتاحية خلال الفترة من 2012م إلى 2016م

ويتضح من الشكل السابق أن أكثر الكلمات المفتاحية ترددًا بخلاف إدارة البيانات البحثية Research Data management هي الكلمات: معالجة البيانات Data Curation، والبيانات البحثية research data، العلم المفتوح open science، وإدارة البيانات data management، والخدمات services، والمكتبات الأكاديمية academic libraries.

ثانيًا: الفترة من 2017م إلى 2019م:

يوضح الشكل التالي مدى تردد الكلمات المفتاحية الواردة في الدراسات المكشفة في قاعدة بيانات ISI WoS خلال الفترة 2017م – 2019م في موضوع إدارة البيانات البحثية:



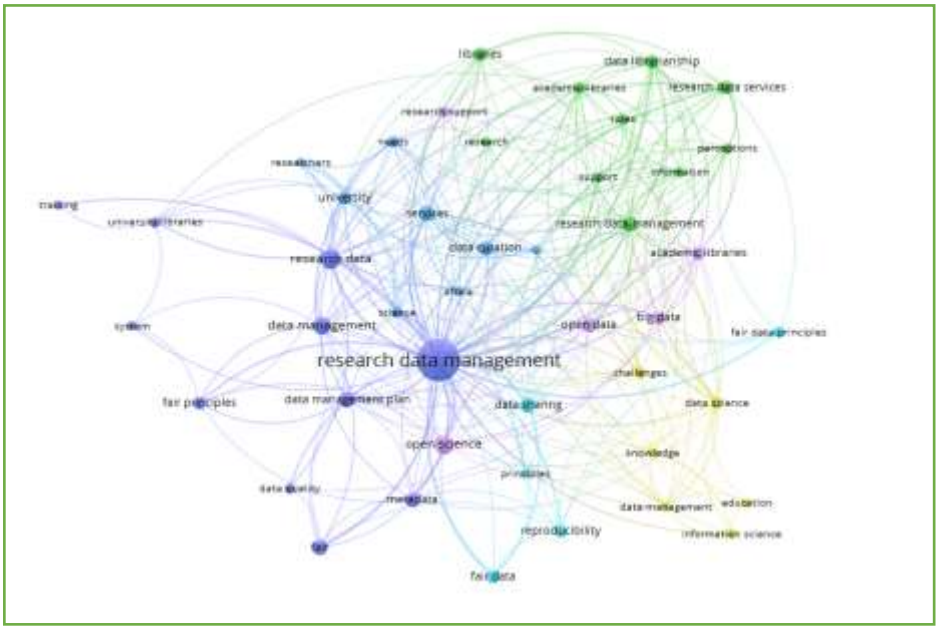
شكل رقم (5) تردد الكلمات المفتاحية خلال الفترة من 2017م إلى 2019م

يتضح من الشكل السابق أن أكثر الكلمات المفتاحية ترددًا بخلاف إدارة البيانات البحثية Research Data management هي الكلمات: البيانات البحثية research data، والعلم المفتوح open science، ومعالجة البيانات data curation، ومشاركة البيانات data sharing، وإدارة البيانات data management، والمكتبات الأكاديمية academic libraries،

والوصول احر open access، والوعي بالبيانات data literacy، والخدمات services، والميتاداتا metadata.

ثالثًا: الفترة من 2020م إلى 2022م:

يوضح الشكل التالي مدى تردد الكلمات المفتاحية الواردة في الدراسات المكشفة في قاعدة بيانات ISI WoS خلال الفترة 2020م – 2022م في موضوع إدارة البيانات البحثية:



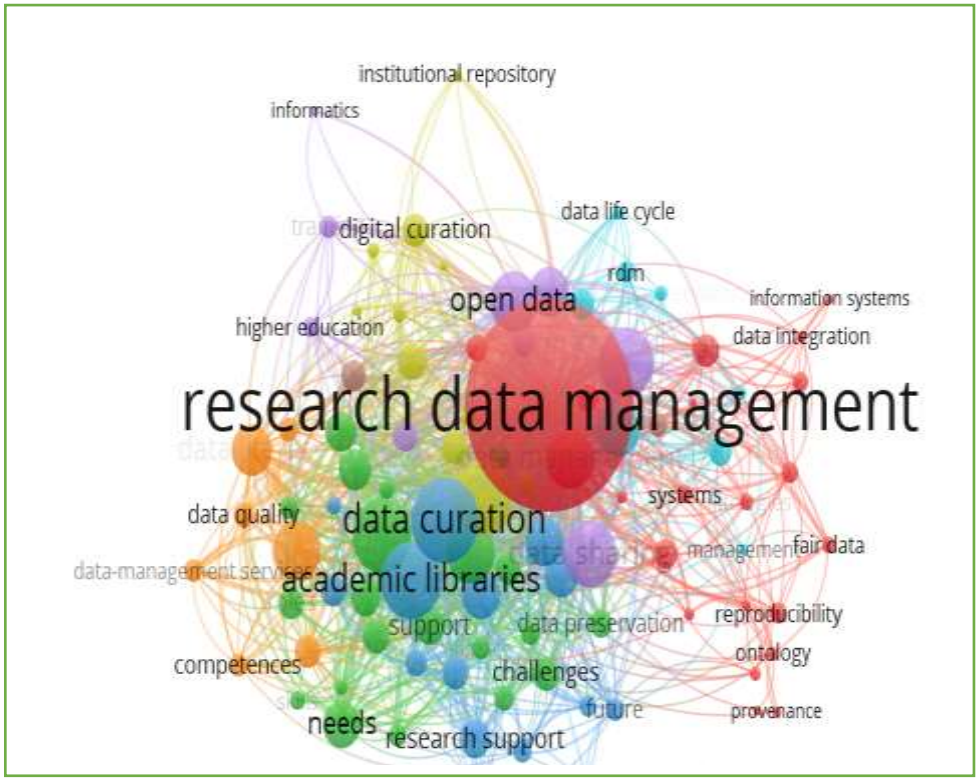
شكل رقم (6) كثافة تردد الكلمات المفتاحية خلال الفترة من 2020م إلى 2022م

يتضح من الشكل السابق أن أكثر الكلمات المفتاحية ترددًا بخلاف إدارة البيانات البحثية Research Data management هي الكلمات: معالجة البيانات data curation، ومستودع البيانات، data repository، والمكتبات الأكاديمية academic libraries، وإدارة البيانات data management، والميتاداتا metadata، والبيانات البحثية research data.

رابعا: الفترة الإجمالية من 2012م إلى 2022م

يوضح الشكل التالي كثافة تردد الكلمات المفتاحية الواردة في الأعمال المنشورة في مجال إدارة البيانات البحثية خلال الفترة من 2012م إلى 2022م، وكذلك قوة العلاقة بين

مصـــطلح إدارة البيانات البحثية وغيره من المصـــطلحات الأخرى الواردة في الكلمات المفتاحية للأعمال المنشورة. كما يوضح الشكل أيضًا الموضوعات الساخنة ذات العلاقة بإدارة البيانات البحثية. مثل البيانات المفتوحة، ومعالجة البيانات، والمكتبات الجامعية، وجودة البيانات، ودعم البحوث...وغيرها



شكل رقم (7) كثافة تردد الكلمات المفتاحية الواردة في الأعمال المنشورة في إدارة البيانات البحثية

ويلخص الجدول التالي بالأرقام مدى تكرار الكلمات المفتاحية في الفترات الزمنية الثلاث، التي ظهرت في الشـــكال الثلاثة الســـابقة أرقام 3، 4، 5، بالإضـــافة إلى الفترة الزمنية الإجمالية من 2012م إلى 2022م، ومع ملاحظة أنه تم تجميع الكلمات المفتاحية التي كتبت باكثر من صـــيغة واحتســـبتها قاعدة البيانات مصـــطلحات مختلفة، مثل data repository

وdata repositories، وكذلك مثل research data management و – research data management.

جدول رقم (9) الكلمات المفتاحية الأكثر تكرارا في الفترات المختلفة

2022 – 2012		2022 -2020 (17) مصطلح		2019 – 2017 (41) مصطلح		2016 – 2012 (39) مصطلح	
التكرار	الكلمة المفتاحية	التكرار	الكلمة المفتاحية	التكرار	الكلمة المفتاحية	التكرار	الكلمة المفتاحية
295	research data management	66	research data management	128	research data management	101	research data management
65	libraries	15	data curation	26	academic libraries	31	fair data principles
51	research data	14	data repository	26	research data	17	research data
43	data curation	13	libraries	22	open science	17	data management
43	data management	10	data management	19	data curation	15	open science
42	fair data principles	9	metadata	19	data sharing	13	academic libraries
37	open science	8	research data	18	data repository	11	services
36	data sharing	8	science	16	data management	10	data management plan
32	data repository	7	data sharing	15	open access	10	data sharing
31	metadata	7	open access	14	data literacy	9	data curation
26	science	7	research support	14	services	9	metadata
25	services	6	research data services	13	metadata	9	open data
22	access open	5	information literacy	12	science	8	data librarianship

2022 – 2012		2022 -2020 (17) مصطلح		2019 – 2017 (41) مصطلح		2016 – 2012 (39) مصطلح	
التكرار	الكلمة المفتاحية	التكرار	الكلمة المفتاحية	التكرار	الكلمة المفتاحية	التكرار	الكلمة المفتاحية
21	data research services	4	bibliometrics	11	open data	8	libraries
20	data open	4	big data	11	fair principles	8	research data services
17	big data	4	information	10	scholarly communication	8	university
16	university	4	research	8	data services	7	big data
15	information			8	needs	7	reproducibility
14	literacy data			8	support	6	challenges
10	scholarly communication			8	university	6	data science
9	information literacy			7	collaboration	6	needs
				7	research data services	6	science
				6	big data	5	data quality
				6	data preservation	5	information
				6	information	5	perceptions
				6	roles	5	research support
				5	challenges	5	roles
				5	digital curation	5	support
				5	knowledge	5	university libraries

2016 – 2012 (39) مصطلح		2019 – 2017 (41) مصطلح		2022 -2020 (17) مصطلح		2022 – 2012	
الكلمة المفتاحية	التكرار	الكلمة المفتاحية	التكرار	الكلمة المفتاحية	التكرار	الكلمة المفتاحية	التكرار
africa	4	librarians	5				
education	4	perceptions	5				
higher education	4	data	4				
information science	4	data analysis	4				
knowledge	4	data citation	4				
principles	4	data life cycle	4				
research	4	data quality	4				
researchers	4	data-management services	4				
system	4	digital humanities	4				
training	4	ethics	4				
		information literacy	4				
		trends	4				

يتضح من الجدول السابق أنه باستثناء الكلمة المفتاحية research data management فإن: الفترة الزمنية الأولى من 2012م إلى 2016م كانت الكلمة المفتاحية الأكثر تكرارًا في الدراسات المكشفة بقاعدة البيانات ISI WoS هي كلمة fair data principles وهي أربعة مبادئ للبيانات هي أنه (يمكن العثور عليهاFindable ، ويمكن الوصول إليهاAccessible ، وقابلة للتشغيل المتداخلInteroperable ، وقابلة لإعادة الاستخدام Reusable. وقد اقترحها اتحاد العلماء والمنظمات عام 2016 لدعم إعادة استخدام الأصول الرقمية. ومنذ ذلك الحين

يتم تبني هذه المبادئ من قبل مؤسسات البحث في جميع أنحاء العالم[25]) وقد تكررت في 31 وثيقة. ثم جاءت كلمتي research data وdata management وتكررت كل منهما 17 مرة.

الفترة الزمنية الثانية من 2017م إلى 2019م كانت الكلمة المفتاحية الأكثر تكرارًا في الدراسات المكشفة بقاعدة البيانات ISI WoS كلمتان هما: academic libraries و research data وتكررت كل منهما 22 مرة، ثم الكلمة المفتاحية open science وتكررت في الدراسات 17 مرة. وظهرت في هذه الفترة بعض الموضوعات ذات الصلة بإدارة البيانات البحثية مثل الإنسانيات الرقمية، والأخلاقيات، وتحليل البيانات، ودور اختصاصيي المعلومات.

وفي الفترة الزمنية الثالثة من 2020م إلى 2022م كانت الكلمة المفتاحية الأكثر تكرارًا هي الكلمة المفتاحية data curation وتكررت 15 مرة، ثم كلمة data repository وتكررت 14 مرة ثم كلمة libraries وتكررت 13 مرة.

وإجمالا كانت الكلمات الأكثر تكرارًا هي libraries، وdata research وdata curation و data management وfair data principles وقد ظهرت هذه الكلمات في الفترات الزمنية السابقة. كما يظهر كثافة تكرار الكلمة المفتاحية إدارة البيانات البحثية، وعلاقتها القوية بكل الكلمات المفتاحية الأخرى.

4) نتائج الدراسة:

استخدمت هذه الدراسة أسلوب التحليل الببليومتري والتحليل المرئي لدراسة وتحليل منشورات "إدارة البيانات البحثية" المدرجة في قاعدة بيانات ISI WoS، وقامت بتحليل 594 وثيقة مسترجعة، وجاءت النتائج على النحو التالي:

1- ظهرت أول دراسة عن إدارة البيانات البحثية عام 1974م، وشهد عام 2019م أعلى معدل نشر بواقع 88 عملا.
2- أكثر أشكال مصادر المعلومات المدرجة في قاعدة البيانات WOS في موضوع إدارة البيانات البحثية هي مقالات الدوريات ووصل عددها إلى 379 مقالة. يليها أعمال المؤتمرات ثم المراجعات العلمية.
3- تنوعت المجالات الموضوعية التي نشر فيها أعمال في "إدارة البيانات البحثية" وجاء مجال المكتبات والمعلومات في المرتبة الأولى بواقع 310 عملاً، ثم مجال علم الحاسب بواقع 221 عملاً

4- توزعت الأعمال المدرجة في قاعدة البيانات WoS في مجال إدارة البيانات البحثية على 73 دولة من دول العالم جاء في مقدمتها الولايات المتحدة الأمريكية بعدد 138 عملا، ثم ألمانيا بعدد 125 عملاً، ووصل عدد الأعمال المنتمية للدول العربية ثمانية أعمال.

5- تصدرت اللغة الإنجليزية لغات النشر الأعمال المنشورة والمدرجة في قاعدة البيانات WoS في "إدارة البيانات البحثية" بنسبة 89.56 %، يليها اللغة اللمانية ثم البرتغالية.

6- جاء المؤلف البرتغالي .João Rocha da Silva أعلى المؤلفين إنتاجية بعدد 15 عملاً، كما جاءت دورية LECTURE NOTES IN COMPUTER SCIENCE أعلى الدوريات إنتاجية بواقع 25 عملاً في مجال إدارة البيانات البحثية.

7- جاءت المقالة بعنوان Panorama: A Targeted Proteomics Knowledge Base المنشورة عام 2014م أعلى المقالات استشهادًا بها، حيث استشهد بها 139 عملا مدرجا في قاعدة البيانات.

8- اعتمادًا على شبكات الكلمات المفتاحية التي توضح المصطلحات الأكثر ترددًا بالإضافة إلى المصطلح الأساس إدارة البيانات البحثية Research Data Management جاءت المصطلحات التالية: libraries، وresearch data وdata curation و data management وfair data principles . كتخصصات ذات علاقة، وكموضوعات ساخنة في الوقت الحالي في الإنتاج الفكري المكشف في قاعدة البيانات WoS.

الهوامش:

[1] National Science Foundation. (2021).Dissemination and sharing of research results—NSF data management plan requirements. https://www.nsf.gov/bfa/dias/policy/dmp.jsp

2 Bossaller, J., & Million, A. J. (2022). The research data life cycle, legacy data, and dilemmas in research data management. Journal of the Association for Information Science and Technology,1–6. https://doi.org/10.1002/asi.24645

[3] **OpenDOAR. OpenDOAR Statistics, available at: https://v2.sherpa.ac.uk/view/repository_visualisations/1.html**

[4] Cox, A. M., & Pinfield, S. (2016). Research data management and libraries: Current activities and future priorities. Journal of Librarianship and Information Science, 46(4), p.300 https://doi.org/10.1177/0961000613492542

[5] Hu, J., & Zhang, Y. (2017). Discovering the interdisciplinary nature of Big Data research through social network analysis and visualization. Scientometrics, 112,p. 94. https://link.springer.com/content/pdf/10.1007/s11192-017-2383-1.pdf

[6] Hu, J., & Zhang, Y. (2017). Op.Cit.p.p. 91-109.

[7] Xu, Z. & Yu, D. (2019). A Bibliometrics Analysis on Big Data Research (2009–2018). Journal of Data, Information and Management, 3–15. https://doi.org/10.1007/s42488-019-00001-2

[8] Tupan, Kamaludin (2020). PUBLICATION OF RESEARCH DATA MANAGEMENT IN OPEN ACCESS JOURNAL ANALYSIS BASED ON SCOPUS DATA. BACA: Jurnal Dokumentasi dan Informasi, 42 (2) Desember. Pp 215- 226. DOI: https://dx.doi.org/10.14203/j.baca.v41i2.701

[9] PICCOLO, D. M.; TADINI, A. V. W.; TEIXEIRA, H. D.; BOTEGA, L. C.; SANT'ANA, R. C. G.; SANTAREM SEGUNDO, J. E.; ALVES, R. C. V.(2021.. Data quality in research data management: a bibliometric study. Em Questão, Porto Alegre, v. 28, n. 1, p. 159–184,. DOI: 10.19132/1808-5245281.159-184 Available at: https://seer.ufrgs.br/index.php/EmQuestao/article/view/112240

[10] Omwoyo Bosire Onyancha (2016). Open Research Data in Sub-Saharan Africa: A Bibliometric Study Using the Data Citation Index. PUBLISHING RESEARCH QUARTERLY, 32 (3) , pp.227-246 https://081027kq4-1106-y-https-link-springer-com.mplbci.ekb.eg/content/pdf/10.1007/s12109-016-9463-6.pdf

[11] محمود شريف زكريا (2020) إتاحة البيانات البحثية الأولية في قطاعي العلوم البحتة والتطبيقية: دراسة تحليلية على عينة من الدوريات المصرية المكشفة في قاعدة بيانات (ISI Web of Science). المجلة المصرية للمكتبات والوثائق والمعلومات. مج2، ع3 (يناير) ص ص 131- 169

[12] سهير عبد الباسط عيد (2016) الدوريات المصرية في قواعد بيانات الاستشهادات المرجعية العالمية: دراسة تحليلية. المجلة الدولية لعلوم المكتبات والمعلومات. مج3، ع3 (يوليو – سبتمبر) ص 12

[13] Thomson Reuters (2014) Web of Science: brochure.8p. Available at: http://thomsonreuters.com/content/dam/openweb/documents/pdf/scholarly-scientific-research/fact-sheet/wos-next-gen-brochure.pdf

[14] **Web of Science Core Collection. Available at: https://clarivate.com/webofsciencegroup/solutions/web-of-science-core-collection/**

[15] الكشّاف العربي للاستشهادات المرجعية أول كشّاف استشهادات للمنطقة العربية على منصةWeb of Science . [تاريخ الزيارة 2022/4/23] متاح عبر الرابط:
https://clarivate.com/ar/campaigns/first-citation-index-for-the-arab-region /

[16] Payal, Shipra Awasthi and Manorama Tripathi (2019)A Selective Review of Literature on Research Data Management in Academic Libraries. Journal of Library & Information Technology, Vol. 39, No. 6, Nov. p. 338, 339 . DOI : 10.14429/djlit.39.6.14451

[17] Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A.U., Wu, L., Read, E., Manoff, M. & Frame, M. (2011). Data Sharing by Scientists: Practices and Perceptions.. PLoS ONE, 6(6), 1–21. https://doi.org/10.1371/journal.pone.0021101.

[18] محمود شريف زكريا (2020). مصدر سابق، ص. 143

[19] Omwoyo Bosire Onyancha (2016). Open Research Data in Sub-Saharan Africa: A Bibliometric Study Using the Data Citation Index| PUBLISHING RESEARCH QUARTERLY 32 (3). P.232

[20] أميرة محمد سيد، سهير عبد الباسط عيد، أمال طه محمد (2021) البيانات البحثية ومستودعاتها في الجامعات: دراسة نظرية. المجلة المصرية لعلوم المعلومات. مج8، ع2، أكتوبر. ص 440

[21] Alicia Arias-Coello , Clara Simon-Blas, Pablo Arranz-Val, Jose Simon-Martin (2018). Research Data Management in Three Spanish Universities[pp. 195–204].in: Kurbanoğlu et al. (Eds.): ECIL 2017, CCIS 810.p. 195. https://doi.org/10.1007/978-3-319-74334-9_21

[22] Angus, W. & Tedds, J. (2011). Making the Case for Research Data Management Introduction – Doing More with Less Funders' Data Policies. JISC, September

[23] Andreas Heuer (2020). Research Data Management. Information Technology, 62(1). P. 1
https://doi.org/10.1515/itit-2020-0002

[24] أميرة محمد سيد، سهير عبد الباسط عيد، أمال طه. مصدر سابق. ص 441، 442.

[25] **Columbia University(Apr 19, 2022). What are the FAIR Data Principles? Available at: https://library.cumc.columbia.edu/insight/what-are-fair-data-principles**

Data Research Management in the ISI Web of Science database: an analytical study using bibliometric and visual analysis

Dr. Soheir Abdel Baset Eid
Department of Information Sciences
Faculty of Arts - Beni Suef University

Abstract:

This study aims to monitor and analyze Literature in the field of "research data management" Indexed in the ISI Web of Science database. The Study used Bibliometric methods to study the different distributions of this Literature. It also used VOS viewer to visualize and analyze data in order to verify the co-occurrence of keywords in Literature units. The study showed that the year 2019 witnessed the highest publication rate of 88 works, and that journal articles are the most common form of information sources, with a number of 379 articles. The field of library and information Science ranked first among the subject areas published in the research data management with 310 works.

In terms of publishing places, the United States of America came at the fore with 138 works, the English language topped the list with 89.56%, and the Portuguese author João Rocha da Silva, the authors are most productive with 15 works, and the article is titled Panorama: A Targeted Proteomics Knowledge Base, published in 2014, top Articles Cited, based on keyword networks that show the most frequent terms as well as the base term Research Data Management The terms: library, research data, data processing, and FAIR Principles for data are related topics to the term research data management.

Keywords: Research data management; Bibliometric analysis; Visual analysis; Keyword network.