البصمة المعرفية للدوريات العربية في مجال المكتبات والمعلومات: دراسةٌ قائمةٌ على ذخيرةٍ نصيةٍ

أيمن الدكروري
اختصاصي المكتبات
الجامعة الأمريكية بالقاهرة
a.eldakroury@aucegypt.edu

مستخلص:

تنصب هذه الدراسة على الكشف عن البصمة المعرفية للإنتاج الفكري من خلال بناء وتكشيف وتحليل ذخيرةٍ نصيةٍ للدوريات العربية في مجال المكتبات والمعلومات. وفي ضوء ذلك اعتمدت الدراسة على المنهج المسحي، فضلًا عن استعانتها بذخيرةٍ نصيةٍ تم إنشاؤها خصيصًا كي تكون أداةً رئيسةً للدراسة، بالإضافة إلى استخدام برنامج أدوات وورد سميث WordSmith Tools لتحليل البيانات. وقد انضوى هذا التحليل على التحقق من أكثر الكلمات مفتاحيةً keyness، وتكرارًا، وتشتتًا، وكذلك التحقق من الأشكال اللغوية المختلفة للكلمات المفتاحية، وتحليل البيئة اللغوية للكلمات المفتاحية في ضوء تحليل المصاحبة اللغوية collocation، والقوالب اللغوية patterns، على النحو الذي يمكن معه التثبت من المداخل الكشفية، والواصفات، ورؤوس الموضوعات التي تمثل مجال المكتبات والمعلومات بطريقةٍ آليةٍ تعتمد على نصوصٍ واقعيةٍ authentic تعكس الاستخدام الفعلي للمصطلحات عند أهل التخصص.

كلمات مفتاحية: البصمة المعرفية، معالجة اللغة الطبيعية، تحليل المحتوى، تكشيف النصوص، الذخائر النصية، نظم استرجاع المعلومات، الدوريات العربية، مجال المكتبات والمعلومات.

التمهيد:

أتاحت معظم نظم المعالجة الآلية للغة الطبيعية إمكانية الاستفادة من المفاهيم المعلوماتية، والقواعد والظواهر اللغوية، والأدوات الرياضية والأساليب الإحصائية، في سياقٍ متكاملٍ من العلاقات التي كفلت استحداث أنماطٍ غير معهودةٍ في تحليل المحتوى. الأمر الذي أسهم في تأسيس نظريات قامت بدورها بتطوير خوارزميات تحليل النصوص، إلى أن أصبحت أسلوبًا إجرائيًّا لبناء تطبيقاتٍ وأدواتٍ آلية. وقد فتح هذا المجال البابَ على مصراعيه للباحثين في سوق معالجة اللغة الطبيعية لإدخال قواعد معلوماتية- لغوية- منطقية إلى الحاسب الآلي، على النحو الذي يمكن معه استرجاع مخرجاتٍ يمكن معها التحقق من السمات المعلوماتية informational features، والخصائص اللغوية lingual characteristics للنصوص. وهو ما أسماه الباحث في هذه الدراسة بـ "البصمة المعرفية".

حيث يتم التحقق من السمات المعلوماتية في ضوء تكشيف النصوص، وتحليل الكلمات المفتاحية، بينما يتم التحقق من الخصائص اللغوية في ضوء تلازم المصطلحات والقوالب اللغوية.

ولا يمكن لهذا الطرح أن يتم دون الحصول على رصيدٍ ضخمٍ من النصوص المكتوبة أو المنطوقة التي يتم اختزانها على الحاسب الآلي، وهو ما يُعرف بـ "الذخائر النصية". وتُعد الوثائق المنشورة في أحد فروع المعرفة مصدرًا رئيسًا للذخيرة النصية، حيث تؤخذ الكلمة أو الكلمات من العناوين، أو المستخلصات، أو الفقرات، أو النص الكامل؛ تمهيدًا لما سيتم عليها من تحليلٍ كميٍ ونوعيٍ (Schneider, 2004, p 27)، وفي هذه المرحلة يمكن طرْح الذخائر النصية كمنتجٍ يتم الاستفادة منه في مجالاتٍ عدةٍ، كلٌ وفق احتياجاته.

وقد تأتي مرحلةٌ لاحقةٌ يتم فيها تشغيل بعض البرمجيات على البيانات الخام التي تحويها الذخائر النصية، حيث يتم عمل مجموعةٍ من التحليلات الإحصائية النظمية والدلالية على هذه البيانات؛ بهدف فحص كل مصطلحٍ أو كلمةٍ مفتاحيةٍ من حيث تركيبتها المعلوماتية واللغوية، وما يرتبط بها من كلماتٍ تسبقها أو تلحقها. الأمر الذي يسهم بشكلٍ كبيرٍ في التحقق من "البصمة المعرفية" لأي نوعٍ أو شكلٍ من النصوص، فضلًا عن إسهامه في الصياغة المُثلى للمصطلحات، ورؤوس الموضوعات، والواصفات، وعبارات البحث، بشكلٍ مُحْكَمٍ يعكس الاستخدام الفعلي أو الواقعي authentic لها، وبعيدًا تمامًا عن البديهة. ومن ثم يمكن تقديم خدمات المعلومات على النحو الذي يلبي احتياجات المستفيدين بدقةٍ وكفاءةٍ.

مشكلة الدراسة:

إن جودة المعالجة الآلية للغة الطبيعية لا تعزو إلى تمثيلها على الحاسب الآلي فحسب، بل أيضًا إلى مدى الإفادة من خدمات المعلومات، وتطبيقات اللغة في وصف وتحليل النصوص. ولا يمكن بحالٍ من الأحوال تطوير أداء الباحثين ونظم المعالجة دون تطوير، أو ربما ابتكار، منهجياتٍ تتصف بالشمولية والواقعية في آنٍ واحدٍ. ولذا بات من الضروري عند تعاملنا مع النصوص، عمليًّا وبحثيًّا، أن ننظر إليها بشكلٍ متكاملٍ؛ بحيث لا نهتم بجانبٍ ونترك الآخر. فلا نهتم مثلًا بالجانب المعلوماتي فقط، ونترك الجانب اللغوي، أوالعكس؛ لأن الجانبين، في نظر الباحث، بمثابة وجهين لعملةٍ واحدةٍ، فلا يُغفل جانبٌ على حساب آخر.

وعلى الرغم من ذلك فإن أحدًا من الباحثين لم يتطرق لتحليلٍ يَضُم بين طياته السمات المعلوماتية جنبًا إلى جنبٍ مع الخصائص اللغوية. وبالتالي تظل الصورة دائمًا غير واضحةٍ وناقصة المعالم.

أهمية الدراسة:

تكتسب الدراسة أهميتها من أهمية معالجة اللغة الطبيعية، وتحليل المحتوى، وتكشيف النصوص في تلبية احتياجات المستفيدين، وتوفير مقومات الخدمات المعرفية؛ كون التحقق من خصائص وسمات النصوص منهجًا رئيسًا في مجالاتٍ عدةٍ في الحياة العملية: كنظم المعلومات، والترجمة، والقانون، وصناعة المعاجم، والتحليل اللغوي، والتدريس، وغيرها.

كما تكتسب هذه الدراسة أهميتها من كونها تهدف إلى الكشف عن "البصمة المعرفية" للنصوص، على نحوٍ لم يتم التطرق إليه بالبحث والدراسة من قبل في وطننا العربي.

تساؤلات الدراسة:

أ . ما الموضوعات المفتاحية التي تعبر عن مجال المكتبات والمعلومات؟
ب . ما خصائص الكلمات المفتاحية لمجال المكتبات والمعلومات؟
ج . إلى أي مدى تتشتت الكلمات المفتاحية في مجال المكتبات والمعلومات؟
د . ما البيئة اللغوية للكلمات المفتاحية في مجال المكتبات والمعلومات؟

أهداف الدراسة:

انطلاقًا من مشكلة الدراسة وأهميتها، فإن الهدف الرئيس من هذه الدراسة هو محاولة الكشف عن البصمة المعرفية لنصوص مقالات الدوريات في مجال المكتبات والمعلومات، وذلك من خلال:

أ . التحقق من الموضوعات المفتاحية التي تعبر عن مجال المكتبات والمعلومات.
ب. قياس مدى تشتت الكلمات المفتاحية والمصطلحات الكشفية في مجال المكتبات والمعلومات.
ج. تحليل نصوص الإنتاج الفكري لمجال المكتبات والمعلومات في ضوء فحص بيئتها اللغوية.

حدود الدراسة:

أ . الحدود الموضوعية: تنصب هذه الدراسة على تحليل "البصمة المعرفية" للإنتاج الفكري.
ب. الحدود اللغوية: تتناول الدراسة النصوص العربية للإنتاج الفكري في مجال المكتبات والمعلومات.
ج. الحدود النوعية: تتخذ الدراسة من نصوص مقالات الدوريات عينةً ممثِلةً للنصوص العربية الإلكترونية.
د . الحدود الزمنية: تتناول الدراسة نصوص مقالات الدوريات في مجال المكتبات والمعلومات المتوافرة منذ أقدم مقالةٍ إلكترونيةٍ أمكن الحصول عليها؛ أي منذ عام ٢٠٠٠ (العربية ٣٠٠٠، ع١) وحتى نهاية عام ٢٠١٥.

منهج الدراسة وأداة جمع البيانات:

تعتمد الدراسة على المنهج المسحي الذي يضطلع بالمحاولة المنظمة لتحليل وتفسير الوضع القائم والوصول إلى بياناتٍ يمكن تصنيفها وتحليلها (عبد الهادي، ٢٠٠٢، ص ١٠٢)؛ وذلك بهدف تشخيص الواقع تشخيصًا دقيقًا واستقراء الحقائق في محاولةٍ للتحقق من البصمة المعرفية للنصوص اعتمادًا على ذخيرةٍ نصيةٍ كأداةٍ رئيسةٍ في جمع البيانات.

مجتمع الدراسة:

يتمثل مجتمع الدراسة فيما تحويه الذخيرة النصية من نصوصٍ كاملةٍ لمقالات الدوريات في مجال المكتبات والمعلومات التي قام الباحث بحصرها وتجميعها من الشبكة العنكبوتية وقواعد البيانات العربية.

الجدول رقم (١) مكونات الذخيرة النصية لمجال المكتبات والمعلومات

الدورية	عدد المقالات	عدد الكلمات	الحجم بالميجا بايت
١. Cybrarians Journal	١٧٥	٩٠٥١٦	١٦.٧
٢. العربية ٣٠٠٠	٢١٥	٧٦٥٣١	١٠.٢
٣. المجلة الأردنية للمكتبات والمعلومات، (رسالة المكتبة سابقًا)	٧٣	٥٨٨٨٤	٥.٦١
٤. اعلم	٥٤	٣٨٤٤٤	٤.٥٤
٥. دراسات المعلومات	٥٠	٣٥٢٧٧	٦.٠٧
٦. دراسات عربية في المكتبات وعلم المعلومات	٤٥	٣٣٧٣١	٤.٩٢
٧. مكتبات. نت	٦٢	٢٤١٠٢	١.٨٤
الإجمالي	٦٧٤	٣٥٧٤٨٥	٤٩.٨٨

أداة تحليل البيانات:

وتتخذ الدراسة من برنامج وورد سميث WordSmith (Scott, ٢٠١٦) أداةً رئيسةً في القيام بالعمليات الإحصائية التي تشمل: تقدير الدلالة الإحصائية، وطريقة احتمالات سجل الأداء loglikelihood، ومعدل التشتت؛ وذلك بهدف تحليل الذخيرة النصية لمجال المكتبات والمعلومات، ومن ثم التحقق من الموضوعات الأساس والكلمات المفتاحية التي تعبر عن هذا المجال.

وبرنامج أداوت وورد سميث WordSmith Tools عبارةٌ عن حزمة برمجياتٍ مدفوعة الأجر تُستخدَم في تحليل النصوص المتاحة بلغاتٍ عدةٍ، ومنها النصوص العربية. ويقوم على هذا البرنامج مايك سكوت MikeScott الباحث بجامعة ليفربول University of Liverpool وبالتعاون من جامعة أكسفورد OxfordUniversity (Scott, ٢٠١٦). ويتوافر برنامج وورد سميث على تقديم ثلاث عملياتٍ فنيةٍ رئيسةٍ متمثلةٍ في ثلاثة نظمٍ فرعيةٍ، هي: كشاف الكلمات المفتاحية في السياق "كونكورد Concord"، وقائمة الكلمات "وورد لِيست WordList"، ومُحلل الكلمات المفتاحية "كي وورد KeyWord".

وإضافةً إلى هذه النظم الفرعية الثلاثة، فإن برنامج وورد سميث يتيح مجموعةً أخرى من الإمكانات التي تضفي قدرًا أدق من التحليل على الذخائر النصية، ومنها: تشتت الكلمات المفتاحية keyworddispersion، وتحليل تلازم المصطلحات/المصاحبة اللغوية collocation، والكشف عن

القوالب اللغوية patterns (Scott, ٢٠١٦).

إجراءات تحليل البيانات:

استثمرت الدراسة كافة العمليات الفنية التي يتوافر عليها برنامج وورد سميث، فيما عدا النظام الفرعي لقائمة الكلمات؛ إذ لم تكن هناك حاجة لاستخدامه سوى في عمل الذخيرة النصية المرجعية فقط. فما يعني الدراسة هو تكشيف الذخيرة النصية لمجال المكتبات والمعلومات، ومن ثم استرجاع المداخل الكشفية وتحليلها.

وقد رافق هذا التحليل بعض العمليات الفنية التحليلية الأخرى التي ساهمت كثيرًا في التحقق من "البصمة المعرفية" للمداخل الكشفية. وهذه العمليات هي: تكشيف الكلمات المفتاحية في السياق، وتشتت الكلمات المفتاحية، والأشكال الأخرى للكلمات المفتاحية، وتحليل المصاحبة اللغوية للكلمات المفتاحية، وتحليل القوالب اللغوية للكلمات المفتاحية.

التكشيف الآلي للذخيرة النصية:

تم تشغيل برنامج وورد سميث على الذخيرة النصية بغرض إجراء تكشيفٍ آلي لنصوصها. ومن ثم إمكانية التحقق من الكلمات المفتاحية التي تعبّر عن مجال المكتبات والمعلومات.

تحديد/أخذ العينة:

يعطي برنامج وورد سميث عددًا كبيرًا من الكلمات المفتاحية للنصوص التي يتم تكشيفها، فضلًا عن أنه يُمكِّن الباحثين من تحديد عدد المداخل المسترجعة وفقًا للطلب. وقد اكتفت هذه الدراسة بأول عشر كلماتٍ مفتاحيةٍ الأعلى في درجة مفتاحيتها؛ تجنبًا لأي إسهابٍ في الكلمات التي يتم عليها التحليل من شأنه أن ينعكس سلبًا على حجم الدراسة ذاتها.

KeyWords

File Edit View Compute Settings Windows Help

N	Key word	.Freq	%	.RC. Freq	% .RC	Keyness
1	المعلومات	33,777	0.89	117	0.02	8,696.51
2	المكتبات	22,445	0.59	21		6,244.71
3	المكتبة	11,844	0.31	54		2,949.22
4	البحث	14,040	0.37	165	0.03	2,921.94
5	الدراسة	12,817	0.34	109	0.02	2,883.60
6	الإلكترونية	7,578	0.20	16		2,024.06
7	البيانات	7,775	0.21	39		1,911.23
8	الرقمية	6,385	0.17	4		1,793.20
9	والمعلومات	6,712	0.18	15		1,785.72
10	الإنترنت	7,997	0.21	128	0.02	1,507.05

الشكل رقم(١) أول عشر كلماتٍ مفتاحيةٍ في الذخيرة النصية لمجال المكتبات والمعلومات كما وردت في برنامج وورد سميث

الدراسات السابقة

من الدراسات الرائدة في مجال معالجة اللغة الطبيعية وتحليل المحتوى اعتمادًا على الذخائر النصية، تلك الدراسة التي قام بها الباحث التايواني تشن Chen(١٩٩٩)؛ بهدف التحقق من موضوعات الوثائق، حيث اقترح الباحث نموذجًا لمساعدة المستفيدين على استرجاع الوثائق على نحوٍ أكثر فعاليةً. وقد اعتمد هذا النموذج على أربعة معايير رئيسةٍ، وهي: أهمية الكلمات، وتردد الكلمات، والمصاحبة بين الكلمات، والمسافة بين الكلمات. وقد طُبق هذا النموذج على مجموعةٍ من النصوص الصينية المتاحة من خلال ذخيرة سينيكا SinicaCorpus المكوَّنة من ٥ مليون كلمةٍ صينيةٍ مَحْشوَّةٍ annotated وموسومةٍ tagged.

وفي عام ٢٠٠٢، أجرى باحث من الجامعة العبرية (Drori, ٢٠٠٢) دراسةً حول التحقق من موضوعات الوثائق المتاحة من خلال المكتبات الرقمية بالاعتماد على ذخيرةٍ نصيةٍ قوامها ٢٠٠ مقالةٍ علميةٍ بنصوصها الكاملة في موضوعين رئيسين، هما: الجغرافيا، ودراسات الأسرة. واستخدمت الدراسة برنامج TextAnalysis في تحليلها للذخيرة النصية. وأوضحت النتائج أن هذا البرنامج يعمل بكفاءةٍ ودقةٍ عاليةٍ في المساعدة في تحديد الكلمات المفتاحية لمحتويات المكتبات الرقمية. الأمر الذي من شأنه أن يدعم عمل المسئولين عن خدمة البحث بالمكتبات الرقمية في

تعيين الكلمات المفتاحية للمواد التي تقتنيها بطريقةٍ آليةٍ عالية الكفاءة.

كما قدَّم الباحث السلوفيني كانك Kanič (٢٠١٣) ورقةً بحثيةً خلال المؤتمر العلمي الدولي للذخائر النصية اقترح فيها مشروعًا لبناء أداةٍ لتنظيم وتقنين مصطلحات المكتبات وعلم المعلومات المتاحة باللغة السلوفينية اعتمادًا على اختزان وتقييم استخدام النصوص العلمية والفنية في المجال. وفي هذا السياق قام الباحث ببناء ذخيرةٍ نصيةٍ للاستعانة بها في تيسير عمل مفوضية مصطلحات المكتبات The Library Terminology Commissionon وفقًا للإطار الذي أقرته جمعية المكتبات السلوفينية. وتكونت هذه الذخيرة النصية من ١٠٣٠٠ مستخلصٍ شكلت نحو نصف مليون كلمةٍ من النصوص المنشورة قبل عام ١٩٩٩، ما بين رسائل ماجستير، وأطروحات دكتوراه، ومنفرداتٍ، ومقالات دوريات. وقد استخدم الباحث برنامج نيفا NEVA في تحليل الذخيرة النصية.

وخلال ورشة العمل الدولية الأولى للمكتبات الرقمية في علم الموسيقى First International Digital Libraries for Musicology Workshop في عام ٢٠١٤، قام باحثان من جامعة إلينوي بإربانا شامبين University of Illinois at Urbana-Champaign، وصاحبهما في ذلك باحثٌ آخر من مكتبة الجامعة (Downie; Dougan& Bhattacharyya, ٢٠١٤)، بدراسةٍ ببليومتريةٍ لمجموعات مكتبة هاثي تراست الرقمية HathiTrust Digital Library (HTDL) شملت توزيعاتٍ تكراريةً بالمساهمين بالمواد داخل هذه المكتبة، وبلغات المجموعات المقتناه بالمكتبة، وبالموضوعات ورؤوس الموضوعات التي تغطيها هذه المجموعات، وبتواريخ نشرها، والأنواع الأدبية لها. وفي القسم الأخير من الدراسة قام الباحثون ببناء ذخيرةٍ نصيةٍ لمقتنيات هذه المكتبة من خلال عمل مسحٍ ضوئيٍ لكافة المواد المكتوبة المتاحة، وعلى النحو الذي يُمكِّن من البحث في النصوص الكاملة لها. وبذلك استطاع الباحثون التحقق من الكلمات المفتاحية لكل مادةٍ، ومن ثم الاستفادة منها في صياغة رؤوس الموضوعات المناسبة. وفي النهاية أكد الباحثون على جدوى تضمين تكشيف الكلمات المفتاحية القائم على ذخيرةٍ نصيةٍ في تيسير استرجاع المواد المكتبية.

النتائج والتحليل

الكلمات المفتاحية:

يُلاحَظ من الجدول رقم (٢) أن الكلمة المفتاحية "المعلومات" حظيت بأعلى درجة مفتاحية keyness (١١٣٢٥.٢٦)، كما جاءت في الوقت نفسه كأكثر الكلمات المفتاحية تكرارًا (٤٧٠٦٦) ضمن الذخيرة النصية للمكتبات والمعلومات. بينما حظيت الكلمة المفتاحية "المكتبات" بثاني أعلى درجةِ مفتاحيةٍ (٧٦٢٧.٩٩) وثاني أكثر الكلمات المفتاحية تكرارًا (٢٨٣٨٤). فيما حظيت الكلمة المفتاحية "المكتبة" بثالث أعلى درجةِ مفتاحيةٍ (٢٦٣١.٣٩) وثالث أكثر الكلمات المفتاحية تكرارًا (١٨٩٩٠).

الجدول رقم (٢) التوزيع التكراري للكلمات المفتاحية

الكلمة المفتاحية	# التكرار	٪	# التكرار في ذ م	٪ في ذ م	درجة المفتاحية
المعلومات	٤٧٠٦٦	٠.٨	١٨٧	٠.٠٢	١١٣٢٥.٢٦
المكتبات	٢٨٣٨٤	٠.٥٣	٣٥	ـــــــ	٧٦٢٧.٩٩
المكتبة	١٨٩٩٠	٠.٢٨	٩٤	ـــــــ	٢٦٣١.٣٩
البحث	١٦٥٧٩	٠.٣٣	٢٢٤	٠.٠٣	٢٥٦٣.٧٩
الدراسة	١٧١٧٤	٠.٣	٢٢٦	٠.٠٢	٢٥٢٠.٦٩
الإلكترونية	١٤٧٤٠	٠.١٨	٣٣	ـــــــ	٢٢١٨.٤١
البيانات	١٠٥٠٣	٠.١٨	٤٨	ـــــــ	٢٠٠٤.٦٩
الرقمية	٨٨٣٨	٠.١٥	٩	ـــــــ	١٩١٧.٦٨
الإنترنت	١٠١٠٤	٠.١٩	١٣١	٠.٠٢	١٧١٠.١٠
المعرفة	٨٢٤٦	٠.١٦	١٢٣	٠.٠١	١٥٣١.٤٣

 # التكرار: عدد مرات تردد أو ورود الكلمة في الذخيرة النصية موضوع التحليل والدراسة.

٪ : النسبة المئوية للكلمة في الذخيرة النصية موضوع التحليل والدراسة.

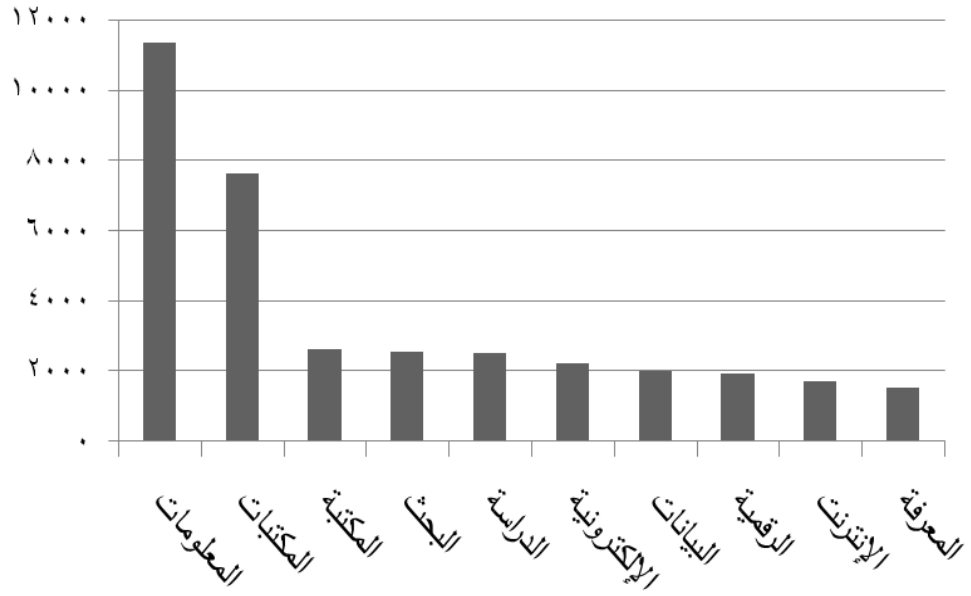
 # التردد في ذ م : عدد مرات تردد أو ورود الكلمة في الذخيرة النصية المرجعية.

٪ في ذ م : النسبة المئوية للكلمة في الذخيرة النصية المرجعية.

كما نالت كل من الكلمة المفتاحية "البحث" و"الدراسة" الترتيب الرابع والخامس تواليًا من حيث درجة المفتاحية (٢٥٦٣.٧٩ و٢٥٢٠.٦٩ تواليًا). بينما حظيت الكلمات المفتاحية: "الإلكترونية" و"الرقمية" و"الإنترنت" بالترتيب السادس، والثامن، والتاسع تواليًا من حيث درجة

المفتاحية (٢٢١٨.٤١ و١٩١٧.٦٨ و١٧١٠.١٠ تواليًا). أما الكلمة المفتاحية "البيانات" والكلمة المفتاحية "المعرفة"، فقد نالتا الترتيب السابع والعاشر تواليًا كأعلى درجةٍ مفتاحيةٍ (٢٠٠٤.٦٩ و١٥٣١.٤٣ تواليًا)، وكأكثر الكلمات المفتاحية تكرارًا (١٠٥٠٣ و٨٢٤٦ تواليًا).

ويوضح الشكل رقم (٢) الكلمات المفتاحية للذخيرة النصية لمجال المكتبات والمعلومات موزعةً تبعًا لدرجة المفتاحية.



الشكل رقم (2) توزيع الكلمات وفقًا لدرجة المفتاحية

الأشكال اللغوية المختلفة للكلمات المفتاحية:

يتعامل برنامج وورد سميث مع الأشكال اللغوية المختلفة للكلمة، والكلمات ذات الأخطاء الإملائية على أنها كلماتٌ مختلفةٌ، فعلى سبيل المثال، نجد أن هناك مجموعةً مختلفةً من السوابق prefixes التي أُلحقت بكلمة "معلومات"، وتتراوح هذه السوابق بين أداة التعريف "ال"، وحرف العطف "واو"، وأداة التعريف "وال"، وحرف الجر "لـ". وقد تعامل وورد سميث مع كل كلمةٍ من هذه الكلمات على أنها كلمةٌ قائمةٌ بذاتها، وليست كلمةً واحدةً.

الجدول رقم (٣) التوزيع التكراري للأشكال اللغوية المختلفة للكلمات المفتاحية

الأسرة اللغوية	الأشكال المختلفة للكلمة المفتاحية					الإجمالي
معلومات	المعلومات	والمعلومات	معلومات	للمعلومات		
	٣٣٧٧٧	٦٧٠٨	٣٤٣٤	٣١٤٧		٤٧٠٦٦
مكتبات	المكتبات	مكتبات	للمكتبات	والمكتبات	بالمكتبات	
	٢٢٤٤٥	٢٣٥٤	١٧٣٨	١١١٣	٧٣٤	٢٨٣٨٤
مكتبة	المكتبة	مكتبة	للمكتبة	بالمكتبة		
	١١٨٤٤	٥٢٠٤	١٠٢٣	٩١٩		١٨٩٩٠
بحث	البحث	بحث	والبحث			
	١٤٠٤٠	١٥٥٧	٩٨٢			١٦٥٧٩
دراسة	الدراسة	دراسة				
	١٢٨١٧	٤٣٥٧				١٧١٧٤
إلكتروني	الإلكترونية	الإلكتروني	الالكترونية	الالكتروني	إلكترونية	
	٧٥٧٨	٣٩١٩	١٣٦٥	١٠٠٨	٨٧٠	١٤٧٤٠
بيانات	البيانات	بيانات				
	٧٧٧٥	٢٧٢٨				١٠٥٠٣
رقمي	الرقمية	الرقمي	رقمية			
	٦٣٨٥	١٧٢٦	٧٢٧			٨٨٣٨
إنترنت	الإنترنت	الانترنت				
	٧٩٩٧	٢١٠٧				١٠١٠٤
معرفة	المعرفة	معرفة				
	٦٦٤١	١٦٠٥				٨٢٤٦

ومن أجل التغلب على هذه المشكلة، فإن البرنامج يتيح إمكانية تجميع الكلمات مع بعضها البعض، بصرف النظر عن أنها أشكالٌ مختلفةٌ للكلمة المفتاحية أم لا. وتظهر هذه الكلمات المجمعة في البرنامج تحت مسمى "الأُسَر اللغوية lemmas". ومن هنا استفاد الباحث من هذه الإمكانية وقام بتجميع الأشكال المختلفة من كل كلمةٍ مفتاحيةٍ من الكلمات موضوع الدراسة، ووضعها في أسرةٍ لغويةٍ واحدةٍ. كما هو موضح في الجدول رقم (٣).

تشتت الكلمات المفتاحية:

يتضح من الجدول رقم (٤) أن أكثر الكلمات المفتاحية تشتتًا هي: المعلومات، والمكتبات، والبيانات؛ وذلك بدرجات تشتت بلغت ٠,٩٦٠، و ٠,٩١٩، و ٠,٩٠٤ تواليًا.

وهذا يعني أن هذه الكلمات الثلاث تشتت أو توزعت على كل المقالات التي تتكون من هذه الذخيرة النصية، وفي كل المواقع تقريبًا، أي أن الكلمة المفتاحية "معلومات"، على سبيل المثال، وردت في كل مقالةٍ من مقالات الذخيرة النصية، فضلًا عن أنها وردت في كل جزءٍ من أجزاء المقالات: العنوان، والمستخلص، والمقدمة، والأجزاء النظرية، والأجزاء العملية، والخاتمة... إلخ.

فيما جاءت الكلمة المفتاحية "المكتبة" في المركز الرابع بدرجة تشتت بلغت ٠,٨٩٩. وجاءت الكلمة المفتاحية "المعرفة" في المركز الخامس بدرجة تشتت بلغت ٠,٨٨٤. أما الكلمة المفتاحية "البحث"، فقد احتلت المركز السادس بدرجة تشتت بلغت ٠,٨٧٥. واحتلت الكلمة المفتاحية "الرقمية" المركز السابع بدرجة تشتت بلغت ٠,٨٥٩. وأتت الكلمة المفتاحية "الإنترنت" في المركز الثامن بدرجة تشتت بلغت ٠,٨٣٨. بينما جاءت الكلمة المفتاحية "الدراسة" في المركز التاسع بدرجة تشتت بلغت ٠,٨٣٣. أما الكلمة المفتاحية "الإلكترونية"، فجاءت متأخرة في المركز العاشر بدرجة تشتت بلغت ٠,٨٠٦.

الجدول رقم (٤) تشتت الكلمات المفتاحية

الكلمة	درجة	الخريطة التخطيطية
المعلومات	٠,٩٦٠	
المكتبات	٠,٩١٩	
المكتبة	٠,٨٩٩	
البحث	٠,٨٧٥	
الدراسة	٠,٨٣٣	
الإلكترونية	٠,٨٠٦	
البيانات	٠,٩٠٤	
الرقمية	٠,٨٥٩	
الإنترنت	٠,٨٣٨	
المعرفة	٠,٨٨٤	

البيئة اللغوية للكلمات المفتاحية:

يتيح برنامج وورد سميث إمكانية إجراء تكشيفٍ للكلمات المفتاحية في السياق KeyWordInContext(KWIC)، بحيث يمكن للمكشفين والمعجميين أن يتحققوا من السياقات المختلفة التي ترِد بها كل كلمةٍ مفتاحيةٍ. وينبثق من هذه العملية الفنية مجموعةٌ أخرى من العمليات الفنية الفرعية التي يتوافر عليها هذا البرنامج، ومن أمثلة هذه العمليات الفنية الفرعية: تحليلُ تلازم المصطلحات/المصاحبة اللغوية collocation، والكشف عن القوالب اللغوية patterns. ويتم في هذا الجزء من الدراسة تحليلُ الكلمات المفتاحية في ضوء هاتين العمليتين الفنيتين.

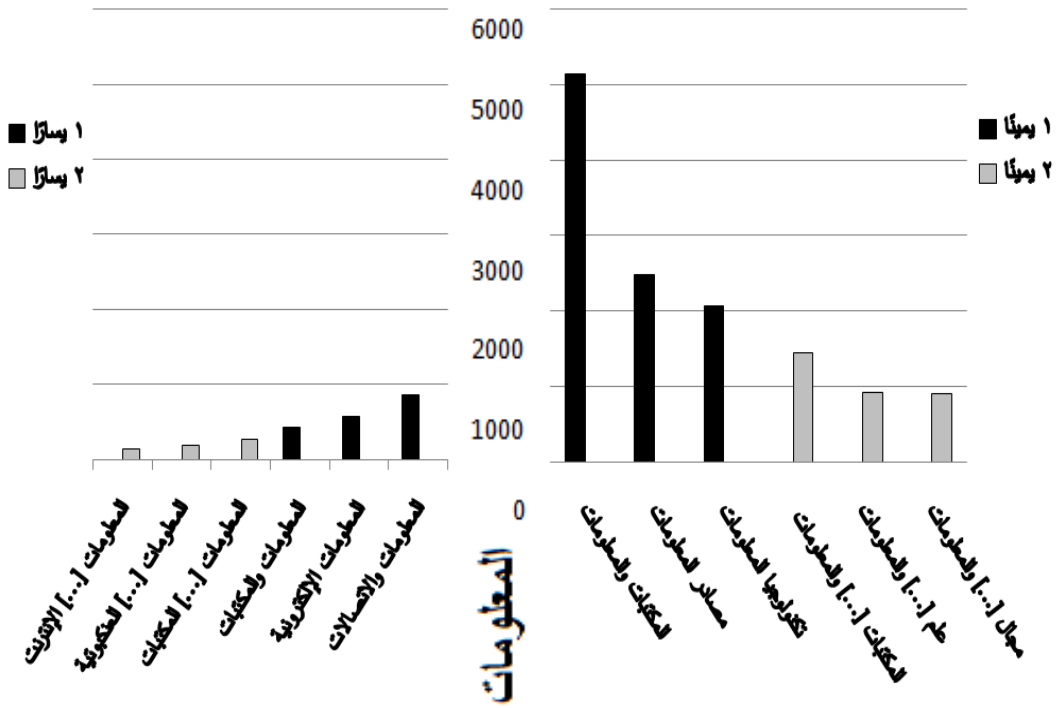
وقد اكتفت الدراسة برصد أكثر كلمتين يمينًا، وأكثر كلمتين يسارًا، وأكثر ثلاث كلمات عمقًا من حيث تكرار ورودها مع الكلمة المفتاحية التي يتم تحليلها.

الكلمة المفتاحية "المعلومات"

يوضح الجدول رقم (٥) والشكل رقم (٣) البيئة اللغوية للكلمة المفتاحية "المعلومات".

الجدول رقم (٥) البيئة اللغوية للكلمة المفتاحية "معلومات"

٢ يمينًا	١ يمينًا	الكلمة المفتاحية	١ يسارًا	٢ يسارًا
المكتبات	المكتبات		والاتصالات	المكتبات
١٤٤٩ واقعة تكرارٍ	٥١٣٥واقعة تكرارٍ		٨٧٤ واقعة تكرارٍ	٢٧٥ واقعة تكرارٍ
علم (+ علوم)	مصادر	معلومات	الإلكترونية	العنكبوتية
٩٢٥واقعة تكرارٍ	٢٤٧٩ واقعة تكرارٍ	(بأشكالها المختلفة)	٥٨٣ واقعة تكرارٍ	٢٠٣ واقعة تكرارٍ
مجال	تكنولوجيا		والمكتبات	الإنترنت
٩٠٨واقعة تكرارٍ	٢٠٧٣ واقعة تكرارٍ		٤٣٦ واقعة تكرارٍ	١٥٠ واقعة تكرارٍ



الشكل رقم (٣) البيئة اللغوية للكلمة المفتاحية "معلومات"

أولًا- المصطلحات المتلازمة/ المتصاحبات اللغوية collocates:

يُلاحَظ من الجدول رقم (٥) والشكل رقم (٣) أن معظم المصطلحات المتلازمة للكلمة المفتاحية "المعلومات" تأتي قبلها، سواءٌ أكانت قبلها مباشرة، كما في هذه الكلمات: "المكتبات" و"مصادر" و"تكنولوجيا"، أم قبلها بكلمةٍ واحدة، كما في هذه الكلمات: "المكتبات" و"علم" و"مجال". فيما كانت الكلمات التي تصاحبت مع الكلمة المفتاحية "المعلومات" بَعدها أقل في عدد مرات تكرارها، وهذه الكلمات هي: "الاتصالات"، و"الإلكترونية"، و"المكتبات".

وبوجهٍ عامٍ، فإن أكثر كلمةٍ متصاحبةٍ مع الكلمة المفتاحية "المعلومات" هي كلمة "المكتبات"، ويتلوها كلمة "مصادر"، ويتلوهما كلمة "تكنولوجيا".

ثانيًا- القوالب اللغوية patterns:

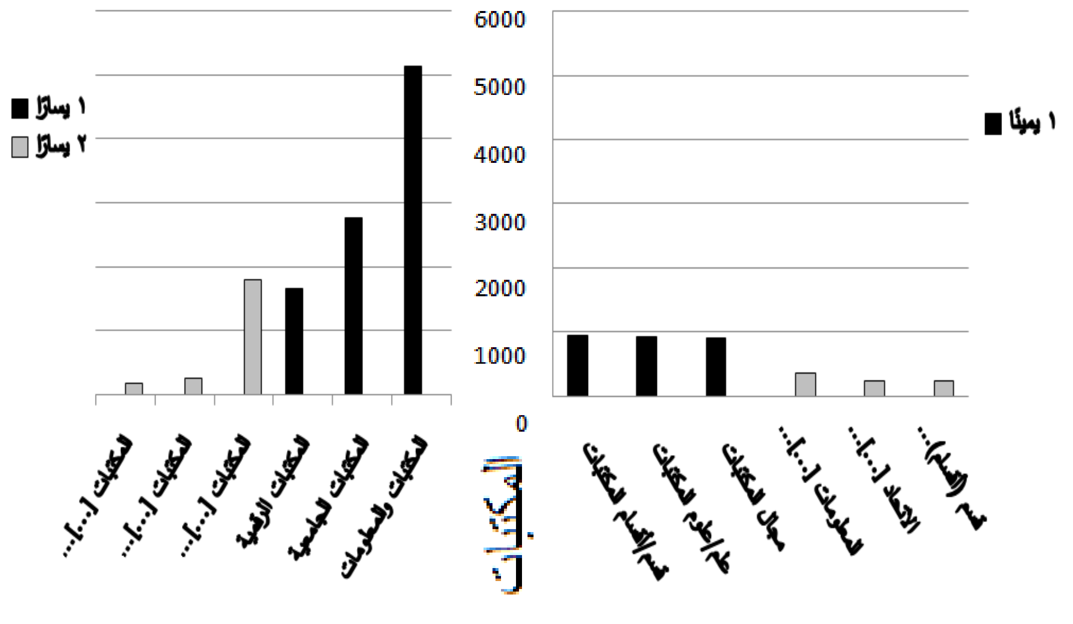
يُلاحَظ من الجدول رقم (٥) والشكل رقم (٣) أن أكثر تركيبٍ لغويٍ وردت به الكلمة المفتاحية "المعلومات" هو تركيب العطف: المكتبات والمعلومات، يتلوه تركيب الإضافة: مصادر المعلومات، وتكنولوجيا المعلومات. فيما جاء تركيب الصفة والموصوف ثالثًا: المعلومات الإلكترونية.

الكلمة المفتاحية "المكتبات"

يوضح الجدول رقم (٦) والشكل رقم (٤) البيئة اللغوية للكلمة المفتاحية "المكتبات".

الجدول رقم (٦) البيئة اللغوية للكلمة المفتاحية "المكتبات"

٢ يمينًا	١ يمينًا	الكلمة المفتاحية	١ يسارًا	٢ يسارًا
المعلومات ٣٦١ واقعة تكرارٍ	قسم (+ أقسام) ٩٤٤ واقعة تكرارٍ	مكتبات (بأشكالها المختلفة)	والمعلومات ٥١٣٥ واقعة تكرارٍ	المعلومات ١٧٩٦ واقعة تكرارٍ
الاتحاد ٢٣٧ واقعة تكرارٍ	علم (+ علوم) ٩٢٥ واقعة تكرارٍ		الجامعية ٢٧٦٩ واقعة تكرارٍ	العربية ٢٥٢ واقعة تكرارٍ
قسم (+ أقسام) ٢٢٤ واقعة تكرارٍ	مجال ٩٠٨ واقعة تكرارٍ		الرقمية ١٦٥٣واقعة تكرارٍ	جامعة ١٧٤ واقعة تكرارٍ



الشكل رقم (٤) البيئة اللغوية للكلمة المفتاحية "المكتبات"

أولًا- المصطلحات المتلازمة:

يُلاحَظ من الجدول رقم (٦) والشكل رقم (٤) أن معظم المصطلحات المتلازمة للكلمة المفتاحية "المكتبات" تأتي بعدها، سواءٌ أكانت بعدها مباشرةً، كما في: "المعلومات"، و"الجامعية"، و"الرقمية"، أم بعدها بكلمة واحدة، كما في كلمة "المعلومات". فيما كانت الكلمات التي تصاحبت مع الكلمة المفتاحية "المكتبات" قبلها أقل في عدد مرات تكرارها، وهذه الكلمات هي: "قسم"، و"علم"، و"مجال".

وبوجهٍ عامٍ، فإن أكثر كلمةٍ متصاحبةٍ مع الكلمة المفتاحية "المكتبات" هي كلمة "المعلومات"، ويتلوها كلمة "الجامعية".

ثانيًا- القوالب اللغوية:

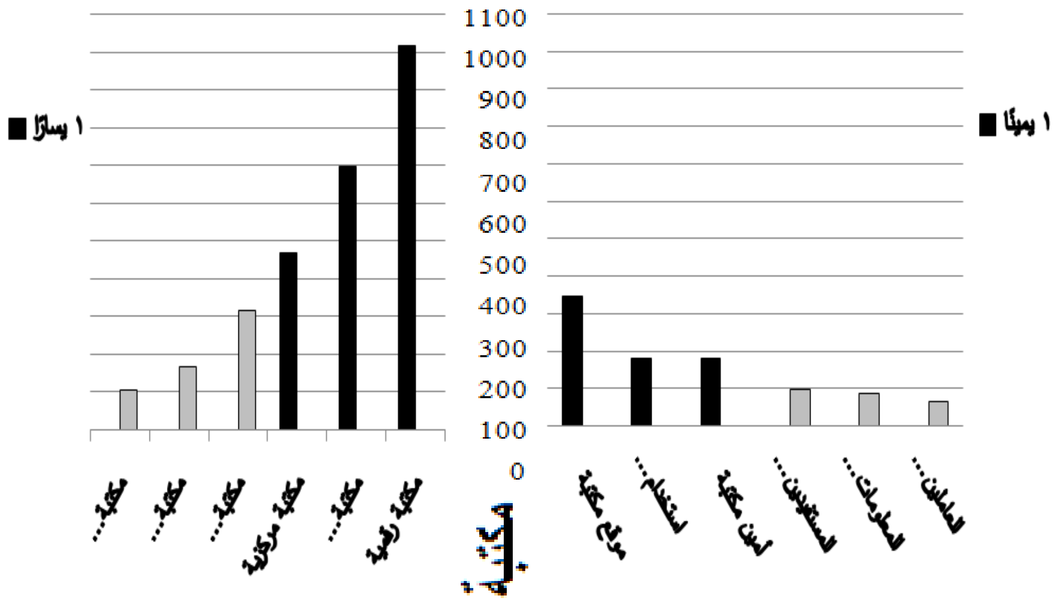
يُلاحَظ من الجدول رقم (٦) والشكل رقم (٤) أن أكثر تركيبٍ لغويٍ وردت به الكلمة المفتاحية "المكتبات" هو تركيب العطف: "المكتبات والمعلومات، يتلوه تركيب الصفة والموصوف: المكتبات الجامعية، والمكتبات الرقمية. فيما جاء تركيب الإضافة ثالثًا: قسم/ أقسام المكتبات، وعلم/علوم المكتبات، ومجال المكتبات.

الكلمة المفتاحية "المكتبة"

يوضح الجدول رقم (٧) والشكل رقم (٥) البيئة اللغوية للكلمة المفتاحية "المكتبة".

الجدول رقم (٧) البيئة اللغوية للكلمة المفتاحية "المكتبة"

٢ يمينًا	١ يمينًا	الكلمة المفتاحية	١ يسارًا	٢ يسارًا
المستفيدون ٩٦	موقع ٣٤٦	مكتبة (بأشكالها المختلفة)	رقمية ١٠١٤	جامعة ٣١٧
المعلومات ٨٤	استخدام ١٧٨		جامعية (+ الأكاديمية) ٦٩٦	الرقمية ١٦٦
العاملون ٦٣	أمين ١٧٨		مركزية ٤٦٩	المعلومات ١٠٦



الشكل رقم (٥) البيئة اللغوية للكلمة المفتاحية "المكتبة"

أولًا- <u>المصطلحات المتلازمة</u>:

يُلاحَظ من الجدول رقم (٧) والشكل رقم (٥) أن معظم المصطلحات المتلازمة للكلمة المفتاحية "المكتبة" تأتي بعدها مباشرةً، كما في: "الرقمية"، و"الجامعية"، و"المركزية". فيما كانت الكلمات التي تصاحبت مع الكلمة المفتاحية "المكتبة" قبلها أقل في عدد مرات تكرارها، وهذه الكلمات هي: "موقع"، و"استخدام"، و"أمين".

وبوجهٍ عامٍ، فإن أكثر كلمةٍ متصاحبةٍ مع الكلمة المفتاحية "المكتبة" هي كلمة "الرقمية"، ويتلوها كلمة "الجامعية".

ثانيًا- <u>القوالب اللغوية</u>:

يُلاحَظ من الجدول رقم (٧) والشكل رقم (٥) أن تركيب الصفة والموصوف هو أكثر تركيبٍ لغويٍ ترِد به الكلمة المفتاحية "مكتبة": مكتبة رقمية، ومكتبة جامعية، ومكتبة مركزية. يتلوه تركيب الإضافة: موقع مكتبة، واستخدام مكتبة، وأمين مكتبة.

الكلمة المفتاحية "البحث"

يوضح الجدول رقم (٨) والشكل رقم (٦) البيئة اللغوية للكلمة المفتاحية "البحث".

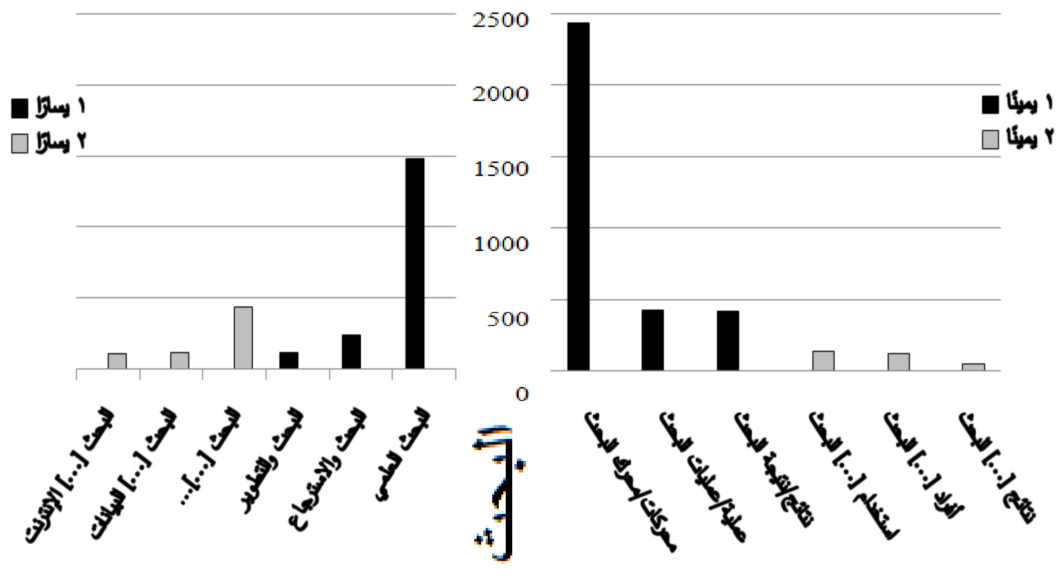
الجدول رقم (٨) البيئة اللغوية للكلمة المفتاحية "البحث"

٢ يمينًا	١ يمينًا	الكلمة المفتاحية	١ يسارًا	٢ يسارًا
استخدام ١٣٣	محركات (+ محرك) ٢٤٣٩	بحث (بأشكالها المختلفة)	العلمي ١٤٧٨	المعلومات ٤٣٦
أفراد ١١٤	عملية (+ عمليات) ٤٢٦		والاسترجاع ٢٣٧	قواعد ١١٤
نتائج ٤٩	نتائج (+ نتيجة) ٤١٤		والتطوير ١١٤	الإنترنت ١٠٩

أولًا- المصطلحات المتلازمة:

يُلاحَظ من الجدول رقم (٨) والشكل رقم (٦) أن معظم المصطلحات المتلازمة للكلمة المفتاحية "البحث" تأتي قبلها مباشرةً، كما في هذه الكلمات: "محرك/محركات"، و"عملية/ عمليات"، و"نتائج/ نتيجة". وفي الوقت نفسه نجد أن الكثير من المصطلحات المتلازمة للكلمة المفتاحية "البحث" تأتي أيضًا بعدها، كما في: "العلمي"، و"الاسترجاع"، و"التطوير".

وبوجهٍ عامٍ، فإن أكثر كلمةٍ متصاحبةٍ مع الكلمة المفتاحية "البحث" هي كلمة "محركات/ محرك"، ويتلوها كلمة "العلمي".



الشكل رقم (٦) البيئة اللغوية للكلمة المفتاحية "البحث"

ثانيًا- القوالب اللغوية:

يُلاحَظ من الجدول رقم (٨) والشكل رقم (٦) أن تركيب الإضافة هو أكثر تركيبٍ لغويٍ ترِد به الكلمة المفتاحية "البحث"؛ وذلك بإجمالي عدد ٣٢٧٩ واقعة تكرارٍ، حيث تشغل كلمة "البحث" موقع المضاف إليه في كل واقعات التكرار هذه، كما في: محرك/ محركات البحث، وعملية/ عمليات البحث، ونتائج/ نتيجة البحث، يتلوه تركيب الصفة والموصوف: البحث العلمي. فيما جاء تركيب العطف ثالثًا: البحث والاسترجاع، والبحث والتطوير.

الكلمة المفتاحية "الدراسة"

يوضح الجدول رقم (٩) والشكل رقم (٧) البيئة اللغوية للكلمة المفتاحية "الدراسة".

الجدول رقم (٩) البيئة اللغوية للكلمة المفتاحية "الدراسة"

٢ يمينًا	١ يمينًا	الكلمة المفتاحية	١ يسارًا	٢ يسارًا
أفراد ٣٤٩	عينة ١٠٤٨	دراسة (بأشكالها المختلفة)	ميدانية ٣٤٢	تحليلية ٧٥
المكتبات ١٤٣	مجتمع ٧٧٧		تحليلية ٢٠١	مقارنة ٣٣
المشاركين ١٤١	نتائج ٥٣٠		حالة ١٩٨	المكتبات ٢١

أولًا- المصطلحات المتلازمة:

يُلاحَظ من الجدول رقم (٩) والشكل رقم (٧) أن معظم المصطلحات المتلازمة للكلمة المفتاحية "الدراسة" تأتي قبلها؛ سواءٌ قبلها مباشرة، كما في هذه الكلمات: "عينة"، و"مجتمع"، و"نتائج". أم قبلها بكلمةٍ واحدةٍ، كما في: "أفراد". فيما كانت الكلمات التي تصاحبت مع الكلمة المفتاحية "الدراسة" قبلها أقل في عدد مرات تكرارها، وهذه الكلمات هي: "ميدانية"، و"تحليلية"، و"حالة".

وبوجهٍ عامٍ، فإن أكثر كلمةٍ متصاحبةٍ مع الكلمة المفتاحية "الدراسة" هي كلمة "عينة"، ويتلوها كلمة "مجتمع"، وكلمة "نتائج".

ثانيًا- القوالب اللغوية:

يُلاحَظ من الجدول رقم (٩) والشكل رقم (٧) أن تركيب الإضافة هو أكثر تركيبٍ لغويٍ ترِد به الكلمة المفتاحية "الدراسة": عينة الدراسة، ومجتمع الدراسة، ونتائج الدراسة، يتلوه تركيب الصفة والموصوف: دراسة ميدانية، ودراسة تحليلية.

النتائج العامة ومناقشتها

تمت مناقشة النتائج وفقًا لما هو متاح من دراساتٍ تبدو على صلةٍ بشكلٍ أو بآخر بموضوع الدراسة الحالية؛ إذ وجِد أن بعض النتائج التي خرجت بها هذه الدراسة تتماس مع نتائج بعض الدراسات الببليومترية.

الموضوعات المفتاحية التي تعبر عن مجال المكتبات والمعلومات:

هناك ستة موضوعاتٍ مفتاحيةٍ تسيطر بشكلٍ أساسٍ على هذا المجال. ويمكن ترتيب هذه الموضوعات المحورية تنازليًا على النحو التالي: المكتبات والمعلومات (٥١٣٥ واقعة تكرارٍ)، والمكتبات الجامعية (٣٤٦٥ واقعة تكرارٍ)، والمكتبات الرقمية (٢٦٦٧ واقعة تكرارٍ)، ومصادر المعلومات (٢٤٧٩ واقعة تكرارٍ)، ومحركات البحث (٢٤٣٩ واقعة تكرارٍ)، وتكنولوجيا المعلومات (٢٠٧٣ واقعة تكرارٍ).

وربما تتفق هذه النتيجة مع النتيجة التي توصل إليها فتحي عبد الهادي في دراسته التي توصلت إلى أن موضوعات المكتبات وعلم المعلومات تستحوذ على النصيب الأكبر من الإنتاج الفكري في المجال، وذلك على حساب الإنتاج الفكري في موضوعات الوثائق والأرشيف (٥٪) (عبد الهادي، ١٩٩٠).

كما تتفق هذه النتيجة مع ما توصلت إليه دراسة أسامة السيد محمود التي أثبتت أن الأرشيف والوثائق من الموضوعات التي يقل فيها حجم الإنتاج الفكري في المجال بشكلٍ عامٍ، وذلك مقارنةً بنظيره في موضوعات المكتبات وعلم المعلومات (السيد، ٢٠٠٠).

وتتفق هذه النتيجة أيضًا مع دراسة علاء عبد الستار مغاوري التي أكدت على أن موضوعات الأرشيف والوثائق العربية أحد الموضوعات التي يقل فيها اهتمام الباحثين العرب غير المصريين في مجال المكتبات وعلم المعلومات (مغاوري، ٢٠٠٤).

وعلى الرغم من أن موضوعات الأرشيف والوثائق في دراسة مها أحمد إبراهيم قد احتلت المرتبة السابعة (٥,٣٪) وليس المرتبة الأخيرة، فإن نسبتها إلى إجمالي حجم الإنتاج الفكري في المجال ككل تظل ضئيلةً (إبراهيم، ٢٠٠٩).

كذلك تتفق هذه النتيجة مع ما توصلت إليه رسالة أيمن الدكروري (٢٠١٣) في أن موضوعات المكتبات وعلم المعلومات تستحوذ على النصيب الأكبر في دراسات التخصص (٩٢,٩٪)، بينما لم تحصد موضوعات الوثائق سوى القليل (٧,١٪).

خصائص الكلمات المفتاحية لمجال المكتبات والمعلومات:

اقتصرت الكلمات المفتاحية التي تعبر عن مجال المكتبات والمعلومات على المصطلحات المكوَّنة من كلمةٍ واحدةٍ، وكلمتين، وثلاث كلماتٍ فقط. بينما لم تكن للكلمات المفتاحية المكوَّنَة من أربع كلماتٍ أو أكثر دلالةٌ إحصائيةٌ، وذلك على النحو التالي:

على مستوى المصطلحات المكوَّنَة من كلمةٍ واحدةٍ:

- أبرز عشر كلماتٍ مفتاحيةٍ يمكن التعبير بها عن الإنتاج الفكري من المقالات هي: المعلومات (٤٧٠٦٦ واقعة تكرارٍ)، و المكتبات (٢٨٣٨٤ واقعة تكرارٍ)، المكتبة (١٨٩٩٠ واقعة تكرارٍ)، والبحث (١٦٥٧٩ واقعة تكرارٍ)، والدراسة (١٧١٧٤واقعة تكرارٍ)، والإلكترونية (١٤٧٤٠ واقعة تكرارٍ)، والبيانات (١٠٥٠٣ واقعة تكرارٍ)، والرقمية (٨٨٣٨ واقعة تكرارٍ)، والإنترنت (١٠١٠٤ واقعة تكرارٍ)، و المعرفة (٨٢٤٦ واقعة تكرارٍ).
- يميل الباحثون إلى استخدام مصطلح "المعلومات" أكثر من مصطلح "المكتبات" في نصوص المقالات، فيما يأتي مصطلح "الوثائق" متأخرًا كثيرًا عنهما. الأمر الذي يعكس إقبال الباحثين في التخصص على الكتابة في هاتين الموضوعين الأولين على حساب الموضوع الأخير.
- يغلب على معظم المقالات في المجال الطابع البحثي والدراسي وليس الطابع الإعلامي وفقًا للذخيرة النصية موضوع الدراسة.

- تحظى الموضوعات التكنولوجية بصفةٍ عامةٍ، ولاسيما تلك المتعلقة بالإنترنت، باهتمامٍ كبيرٍ في كتابات التخصص.

على مستوى المصطلحات المكوَّنَة من كلمتين:

أبرز عشر كلماتٍ مفتاحيةٍ يمكن التعبير بها عن الإنتاج الفكري من المقالات هي: المكتبات والمعلومات (٥١٣٥ واقعة تكرارٍ)، والمكتبات الجامعية (٣٤٦٥ واقعة تكرارٍ)، والمكتبات الرقمية (٢٦٦٧ واقعة تكرارٍ)، و مصادر المعلومات(٢٤٧٩واقعة تكرارٍ)، ومحركات البحث (٢٤٣٩ واقعة تكرارٍ)، وتكنولوجيا المعلومات (٢٠٧٣ واقعة تكرارٍ)، و البحث العلمي (١٤٧٨ واقعة تكرارٍ). وعينة الدراسة (١٠٤٨ واقعة تكرارٍ)، وأقسام المكتبات (٩٤٤ واقعة تكرارٍ)، وعلم المكتبات (٩٢٥ واقعة تكرارٍ).

على مستوى المصطلحات المكوَّنَة من ثلاث كلمات:

أبرز عشر كلماتٍ مفتاحيةٍ يمكن التعبير بها عن الإنتاج الفكري من المقالات هي: المكتبات ومراكز/ وعلم المعلومات (١٧٩٦ واقعة تكرارٍ)، وعلم المكتبات والمعلومات (٩٢٥ واقعة تكرارٍ)، ومجال المكتبات والمعلومات (٩٠٨ واقعة تكرارٍ)، والمكتبات المركزية الجامعية (٣١٧ واقعة تكرارٍ)، والمكتبات الجامعية الرقمية (١٦٦ واقعة تكرارٍ)، وتكنولوجيا المعلومات والاتصالات (٤٥٤ واقعة تكرارٍ)، والبحث عن المعلومات (٤٣٦ واقعة تكرارٍ)، وأفراد مجتمع الدراسة (٣٢٥ واقعة تكرارٍ)، ومصادر المعلومات الإلكترونية (٢٥٨ واقعة تكرارٍ)، وأقسام [المعلومات/ الوثائق] والمكتبات (٢٢٤ واقعة تكرارٍ).

تشتت الكلمات المفتاحية في مجال المكتبات والمعلومات:

اختلفت الكلمات المفتاحية فيما بينها من حيث مدى التشتت؛ إذ جاءت الكلمات المفتاحية: المعلومات، والمكتبات، والبيانات بأعلى درجة تشتتٍ (٠,٩٦٠، و٠,٩١٩، و٠,٩٠٤ تواليًا)، تلاها باقي الكلمات المفتاحية.

البيئة اللغوية للكلمات المفتاحية في مجال المكتبات والمعلومات:

على مستوى المصطلحات المكوَّنَة من كلمتين:

تفضل بعض الكلمات المفتاحية أن تتصاحب مع كلماتٍ على حساب كلماتٍ أخرى، فمثلًا تفضل كلمة "معلومات" أن تتصاحب مع كلمة "مكتبات" وكلمة "مصادر" وكلمة "تكنولوجيا" أكثر من كلمة "وثائق" وكلمة "أوعية" وكلمة "تقنيات" تواليًا. كما تفضل كلمة "المكتبات" أن تتصاحب مع كلمة "الجامعية" وكلمة "الرقمية" أكثر من كلمة "الأكاديمية" وكلمة "الإلكترونية" تواليًا.

على مستوى المصطلحات المكوَّنَة من ثلاث كلمات:

تفضل بعض الكلمات المفتاحية أيضًا أن تتصاحب مع كلماتٍ على حساب كلماتٍ أخرى، فمثلًا تفضل كلمتا "المكتبات" و"المعلومات" أن تتصاحبا مع كلمة "مراكز" وكلمة "علم" أكثر من كلمة "مرافق" وكلمة "مجال" أو "تخصص" تواليًا؛ وذلك فيما يخص الكلمة الواقعة بينهما.

أما فيما يخص الكلمة الواقعة قبلهما، فنجد أنه عادةً ما يكون محجوزًا لكلمة "علم" وكلمة "مجال" أكثر من كلمة "تخصص". ويشير ذلك إلى تفضيل كلمة "المعلومات" لكلمة "علم" دون غيرها؛ أي: علم المعلومات. بينما الوضع أكثر مرونةً مع كلمة "المكتبات" التي قد تتصاحب مع هذه الكلمات جميعًا (علم المكتبات، أو مجال المكتبات، أو تخصص المكتبات).

كما تفضل كلمتا "المعلومات" و"مصادر" أن تتصاحبا مع كلمة "الإلكترونية" أكثر من كلمة "الرقمية"، وفي الوقت نفسه تفضل كلمتا "المعلومات" و"الإلكترونية" أن يتصاحبا مع كلمة "مصادر" أكثر من كلمة "أوعية".

الخلاصة:

حاولت هذه الدراسة التحقق من البصمة المعرفية للدوريات العربية في مجال المكتبات والمعلومات من خلال بناء وتكشيف وتحليل ذخيرةٍ نصيةٍ للمجال. وقد اشتمل ذلك على التحقق من تكشيف، وتكرار، وتشتت الكلمات المفتاحية، وأشكالها اللغوية، وبيئتها اللغوية. وفي هذا السياق أوضحت الدراسة منهجيتها المتمثلة في منهج البحث وأدواته، ومجتمع الدراسة، وطريقة جمع البيانات، وبناء الذخيرة النصية موضوع الدراسة والتحليل، وأدوات تحليل البيانات وإجراءاته، وتحديد العينة، وانتهاءً بالتحليل وعرض النتائج.

هذا، ويمكن أن يفيد التحليل الذي تم في القيام بفحص البصمة المعرفية لأي نوعٍ من أنواع مصادر المعلومات بطريقةٍ آليةٍ تمامًا، على نحوٍ يكفل توفير الوقت والجهد والتكلفة. فعلى سبيل المثال يمكن للقائمين على المكتبات الرقمية العربية الاستعانة بالقالب المنهجي المقترح في هذه الدراسة، إضافةً إلى برنامج وورد سميث، للقيام بتكشيف محتوياتها دون الاستعانة بفريقٍ من المكشفين، وما يتطلبه ذلك من وقتٍ طويلٍ، وتكلفةٍ عاليةٍ، ومجهودٍ كبيرٍ من أجل الاضطلاع بهذه المهمة.

المراجع

أولًا- المراجع العربية:

إبراهيم، مها أحمد. (٢٠٠٩، مايو). الإنتاج الفكري العربي في مجال المكتبات والمعلومات في مطلع القرن الحادي والعشرين؛ دراسة ببليومترية. مجلة مكتبة الملك فهد الوطنية. (١٥)١. ١٧٦-٢٤٩.

الدكروري، أيمن. (٢٠١٣). تكشيف المواقع العربية على العنكبوتية العالمية: دراسة واقعية (إمبريقية) (رسالة ماجستير). جامعة القاهرة.

السيد، أسامة. (٢٠٠٠، يوليو). الإنتاج الفكري المصري في المكتبات والمعلومات ١٨٨٢-١٩٩٥؛ دراسة في السمات والخصائص. الاتجاهات الحديثة في المكتبات والمعلومات. (٧)١٤. ١٥٥-١٧٥.

محمد فتحي عبد الهادي. (١٩٩٠، يناير). دراسة تحليلية للإنتاج الفكري العربي في مجال المكتبات والمعلومات الصادر عام ١٩٨٧م. عالم الكتب، مج ١١، ع ١. ص ص ٢٠-٢٧.

ـــ. (٢٠٠٢). البحث ومناهجه في علم المكتبات والمعلومات. القاهرة: المكتبة الأكاديمية. ٢٥٤ ص.

مغاوري، علاء عبد الستار. (٢٠٠٤، يناير). الإنتاجية العلمية للباحثين العرب غير المصريين في دوريات المكتبات والمعلومات العربية؛ دراسة تحليلية مقارنة. الاتجاهات الحديثة في المكتبات والمعلومات. (١٢)٢١. ١٠٩-١٢٩.

ثانيًا- المراجع الأجنبية:

Chen, K. (1999). Automatic Identification of Subjects for Textual Documents in Digital Libraries. Los Alamos, NM: Los Alamos National Laboratory.

Downie, J. S.; Dougan, K. & Bhattacharyya, S. (2014). The HathiTrust Corpus: A Digital Library for MusicologyResearch? First International Digital Libraries for Musicology Workshop (DLfM 2014), London, UK. Retrieved from http://www.music ir.org/mirex/DLfM_proceedings/long_papers/dlfm2014_submission_33.pdf

Drori, O. (2002). Identifying the Subject of Documents in Digital Libraries Automatically Using .Proceedings of the 3rd International Workshop on New Developments in Digital Libraries. Retrieved from http://leibniz.cs.huji.ac.il/tr/acc/2002/HUJI-CSE-LTR-2002-40_drori062002b.pdf

Kanič, I. (2013). Slovene Specialized Text Corpus of Library and Information Science – an Advanced Lexicographic Tool for Library Terminology Research .Proceedings of the international scientific conference, 2013. Retrieved from http://corpora.phil.spbu.ru/Works2013/Kani%C4%8D.pdf

Schneider, J. W. (2004). Verification of bibliometric methods' applicability for thesaurus construction.(Doctoral dissertation). Aalborg: Department of Information Studies, Royal School of Library and Information Science. p. 27.

Scott, M. (2016).Introduction to Word Smith Tools; version 7.0. Retrieved from http://lexically.net/downloads/version7/HTML/index.html?getting_started.htm