دار المنظومة
DAR ALMANDUMAH
الرواد في قواعد المعلومات العربية

العنوان: محركات بحث الوسائط المتعددة على الشبكة العنكبوتية : المفاهيم، الأدوات، الأنواع
المصدر: مجلة جامعة ناصر الأممية
الناشر: جامعة ناصر الأممية
المؤلف الرئيسي: المدني، محمد الصغير
المجلد/العدد: ع5
محكمة: نعم
التاريخ الميلادي: 2010
الصفحات: 237 - 261
رقم MD: 823573
نوع المحتوى: بحوث ومقالات
قواعد المعلومات: EcoLink, HumanIndex
مواضيع: الإنترنت
رابط: http://search.mandumah.com/Record/823573

© 2021 دار المنظومة. جميع الحقوق محفوظة.
هذه المادة متاحة بناء على الإتفاق الموقع مع أصحاب حقوق النشر، علما أن جميع حقوق النشر محفوظة. يمكنك تحميل أو طباعة هذه المادة للاستخدام الشخصي فقط، ويمنع النسخ أو التحويل أو النشر عبر أي وسيلة (مثل مواقع الانترنت أو البريد الالكتروني) دون تصريح خطي من أصحاب حقوق النشر أو دار المنظومة.

محركات بحث الوسائط المتعددة على الشبكة العنكبوتية
المفاهيم، الأدوات، الأنواع

د. محمد الصغير المدني
جامعة ناصر الأممية

مستخلص

هذه الدراسة تتناول محركات بحث الوسائط المتعددة على شبكة المعلومات الدولية (الإنترنت)، من حيث المفاهيم العامة والأدوات اللازمة والأنواع المختلفة لمحركات البحث. الدراسة قسمت إلى عدة محاور رئيسية، أهمها مفهوم محركات البحث، وكذلك التعرف على محركات بحث الوسائط المتعددة، وأنواعها، بالإضافة إلى مراحل عمل محركات بحث الوسائط المتعددة، كما تم إستعراض بعض نماذج للمحركات الأساسية كمحركات بحث الفيديو، ومحركات بحث الصور، ومحركات بحث الصوت، وصولا إلى دراسة بعض نموذج محركات بحث الوسائط المتعددة ومعرفة مدى فاعليتها. وأخيرا تم وضع بعض النتائج التي خلصت إليها الدراسة.

المقدمة

تعد شبكة المعلومات الدولية (الإنترنت) المكتبة العالمية الفاعلة لكل أنواع المعلومات والمعارف على اختلاف أنواعها، وتعتبر الشبكة العنكبوتية (WWW) قبلة الباحثين عن المعلومة في مختلف فروع المعرفة. لقد مكنت الشبكة العنكبوتية الباحثين من الحصول على مصادر المعلومات في شتى المجالات بكل سهولة، وباللغة التي يرغبها الباحث، وبأشكالها المتعددة المرئية أو المسموعة والنصوص المكتوبة. إن التطور الذي حدث في طرق البحث واقتناء المعلومات من قبل مستخدمي الشبكة

العنكبوتية كان مذهلا، وهو مازال يأخذ في التطور الإيجابي مع ما تفرزه الشبكة العنكبوتية من صقل تلقائي لمحركات البحث لتواكب الطلب المتزايد على مصادر المعلومات. حيث عملت محركات البحث على حصر شامل ومتخصص لمصادر المعلومات المرئية والمسموعة على الشبكة العنكبوتية من خلال صفحات أو مواقع قائمة.

إن تنامى المواقع وصفحات الإنترنت من حيث الكم، جعل المهمة ليست سهلة على الباحثين أن يصلوا إلى مصادر المعلومات بشكل سريع، الأمر الذي جعل الاهتمام بمحركات البحث التي تعمل من خلال أدواتها في فضاء الشبكة العنكبوتية الواسع من الأولويات في عملية البحث السريع والمتزامن حتى تتمكن من حصر كامل لمصادر المعلومات في مجالاتها المختلفة. وتبعا لتلك الزيادة فإن محركات البحث لجأت إلى التخصيص لتخدم كل مجموعة من المحركات موضوعًا أو شكلاً مخصصًا من المعلومات، ولم يقف عند حد التخصص في الشكل والموضوع، إنما تعدى إلى التخصص في مناطق جغرافية محددة تنتمي إليها صفحات العنكبوتية فيما يعرف (Regional Search Engines). كما أخرجت عدة أجيال أكثر تقدما من محركات البحث العادية، كمحركات البحث المتعددة (Meta Search Engines)، مع إضافة بعض النطاقات المحددة مثل مصطلح (Individual) إلى المحركات البحث العادية.

تبحث هذه الدراسة إلى جانب المفاهيم والتعريفات في أحد أنواع محركات البحث؛ وهو محركات بحث الوسائط المتعددة (Multimedia Search engines)؛ كالصوت والصور والحركة والفيديو. من خلال تقسيم محركات البحث إلى قطاعاتها العامة، ثم العمل داخل كل نوع أو قطاع، مع التركيز على شهرة المحركات في هذا القطاع. وإن تناول هذا الموضوع هو زيادة التأكيد على أهمية محركات البحث

المختلفة وزيادة فهم أسرار الإنترنت والشبكة العنكبوتية ورفع ثقافة غير المتخصصين ومهاراتهم في مجال تقنية المعلومات.

الشبكة العنكبوتية العالمية

الشبكة العنكبوتية العالمية بالإنجليزية (World Wide Web)، أو اختصارها (الوب)، هي نظام من مستندات النص الفائق المرتبط بعضها ببعض تعمل من خلال شبكة المعلومات الدولية (الإنترنت). ويستطيع المستخدم تصفّح هذه المستندات باستخدام متصفّح (وب)، كما يستطيع التنقّل بين هذه الصفحات عبر وصلات النص الفائق. وتحوي هذه المستندات على نص صِرْف، أو وسائط متعددة.

تم إنشاؤها سنة 1989م، وكان الغرض منها تيسير الاتصال بين المنظمات الأوربية المجتمعة في سويسرا. والـ(وب) نظام للترابط بين النصوص في الشبكة (الإنترنت). هذا الترابط المتداخل (Hyper linking) بين النصوص في الشبكة يسمح بالانتقال فيما بين الصفحات باستعمال متصفح إنترنت. وصفحات الـ(وب) من الممكن أن تحتوي على نصوص، وصور، وأصوات ومرئيات (فيديو) وصور متحركة، أو برامج تفاعلية وغيرها. ويقوم بتوفير هذه الصفحات على شبكة المعلومات الدولية (الإنترنت) مخدم (وب سيرفر). وهو يقوم بإرسال الصفحات، ومحتوياتها بناء على الطلب من المتصفح نفسه. ومن أهم أدوات شبكة الـ(وب)هي (محركات البحث). وبعد أن أصبحت هناك حاجة لاستعمال خدمات البحث المختلفة لإيجاد المعلومات المطلوبة بسبب ضخامة الشبكة العنكبوتية.

أولا: مفهوم محرك البحث

هو جهاز خادم أو مجموعة من الأجهزة الخادمة المخصصة لتكشيف الصفحات على شبكة الإنترنت، وتقوم بتخزين النتائج وإعداد قوائم بالصفحات المطابقة لاستفسار معين، وتقوم بإعادتها للمستفيد خلال نفس الفترة. وهذه

الكشافات يتم إنتاجها عادة باستخدام الزواحف، كما أن مصطلح محرك البحث غالباً ما يعنى به وصف كل من محركات البحث والأدلة معاً[1].

وقد عرف محمد زايد محركات البحث بأنها: "برمجيات مختصة فى البحث عن المعلومة على شبكة الإنترنت، وتتصل بقواعد البيانات التى تتغذى بصفة مستمرة وتمكن من الحصول على أي معلومة مخزنة على الواب أو على موزع (FTP)"[2].

وعرفها بهاء شاهين بأنها "عبارة عن برامج تجمع البيانات وتعرضها أمام مستخدمى الشبكة فى شكل فهرس بحيث يمكن البحث فيها بيسر وسهولة. وهذه البرامج تعمل فى نظم مستقلة لأجهزة الخدمة المتصلة بدورها بأجهزة خدمة أخرى تتقاسم معها البيانات والمعلومات"[3].

ويورد فهد العبود تعريفاً لمحركات البحث بقوله: "هى أدوات تمكن الباحث من الحصول على المعلومة المطلوبة بسهولة ويسر من بين بلايين الوثائق. وهذه المحركات تقوم بتكشيف صفحات الويب وجمع المعلومات وتنظيمها فى قواعد خاصة. كما أنها تقوم أيضاً بجمع الأخبار والأحداث العالمية ومتابعتها وتقديمها لمستخدميها مجاناً"[4].

ويذكر هشام عزمي تعريفاً لمحرك البحث: "هو برنامج يمكن البحث من خلاله على مصادر المعلومات المتاحة على الإنترنت. وتعتمد محركات البحث على قواعد بيانات ثم تخليقها أوتوماتيكياً، تعمل على تكشيف المصادر المتعلقة بموضوع معين. ويتم ذلك من خلال استخدام عدد من الكلمات الدالة التى تعبر عن موضوع البحث"[5].

وبعد استعراض التعريفات السابقة لمحركات البحث يمكننا القول بأنها برمجيات تساعد المستفيدين على الإبحار بين شبكات المعلومات المرتبط بعضها ببعض باحثين عن المعلومات للاستفادة منها.

ثانيا: مكونات محركات البحث

1. برنامج العنكبوت Spider program
2. برنامج المفهرس Indexer program
3. برنامج محرك البحث Search Engine

1. برنامج العنكبوت

تستخدم محركات البحث برنامج العنكبوت لإيجاد صفحات جديدة على الويب والعمل على إضافتها، ويسمى هذا البرنامج أيضا الزاحف (Crawler) لأنه يقوم بالتحول والبحث في أرجاء الويب المختلفة مطلعا على صفحات الويب التي يزورها ومدققا في محتوياتها بصورة دورية فيأخذ عندها مؤشرات المواقع من عنوان الصفحة والكلمات المفتاحية (KEYWORDS) التي تحتويها، إضافة إلى محتويات محددات الميتا (Meta tags) فيها. ولا تقتصر زيادة برنامج العنكبوت على الصفحة الأولى للموقع بل يتتابع البرنامج لتعقب الروابط (links) الموجودة فيها لزيارة صفحات أخرى. أما الغاية من هذه الزيارات فهي وضع النصوص المنتقاة في نظام الفهارس لمحرك البحث، ليتمكن المحرك من العودة اليها فيما بعد، حيث وضع في الحسبان ضرورة تنظيم محرك البحث بحيث يقوم بزيارات دورية للمواقع المفهرسة لديه للتأكد من التعديلات التي تحدث من حين لآخر على المواقع المسجلة.

2. برنامج المفهرس Indexer Program

يمثل برنامج المفهرس دليلاً أحيانا، من خلال احتوائه على قاعدة بيانات ضخمة تصف صفحات الويب، وتعتمد في التوصيف على المعلومات التي حصلت عليها من برنامج العنكبوت، كما تعتمد بعض المعايير مثل الكلمات الأكثر تكرارا عن غيرها، وتختلف محركات البحث عن بعضها في هذه المعايير، إضافة إلى اختلافها في خوارزميات المطابقة.

3. برنامج محرك البحث Search Engine Program

يبدأ دور برنامج محرك البحث عند كتابة كلمة مفتاحية في مربع البحث (search box) إذ يأخذ هذا البرنامج الكلمة المفتاحية، ويبحث عن صفحات الويب التي تحقق الاستعلام الذي كونه برنامج المفهرس في قاعدة بيانات الفهرس (database index) ثم تعرض نتيجة البحث المتمثلة في صفحات تحتوي وصلات وبعض الشرح لمواقع يعتقد محرك البحث أنها هي ما يبحث عنه المستخدم. وتختلف محركات البحث بعضها بعض عن في أسلوب العمل؛ فمثلا: تحتفظ قاعدة بيانات (Alta vista) بكل تفاصيل صفحة الويب المخزنة، أما غيرها من آليات البحث الأخرى فقد يحتفظ بالعناوين الرئيسية للصفحة فقط، مما يؤدي إلى اختلاف شكل نتائج البحث الظاهرة للمستخدم ودقتها.

ثالثا: أنواع محركات البحث

تختلف محركات البحث في آلية عملها ومضمونها، وهو ما أكد على تصنيفها حسب آلية عملها والمحتوى الذي تقدمه للمستخدم (الباحث). وأهمها: محركات البحث الأجنبية (Foreign Search Engines) التي يمكنها التعامل مع لغات إضافية غير اللغة الإنجليزية، كالفرنسية والإسبانية وغيرها. ومن هذه المحركات، محرك ألتافيستا (AltaVista)، غوغل (Google)، ومحرك إكسايت (Excite). ومحركات البحث المتخصصة (Specialized Search Engines)، يعتبر هذا النوع من محركات البحث ذا أهمية كبيرة خصوصًا على صعيد الأعمال والأبحاث، وحيث إن معظم محتوى الإنترنت باللغة الإنجليزية، ولا تشكل اللغات الأخرى إلا جزءًا بسيطًا من ذلك المحتوى الغزير والوفير، فقد عملت كثير من الشركات على إيجاد محركات بحث تدعم اللغة العربية لإعطاء الفرصة لأكبر عدد من الناس للاستفادة من محتوى الإنترنت الهائل. وسنتطرق هنا لبعض محركات البحث التي تدعم اللغة العربية، وأهمها محرك غوغل الذي أعطى نتائج مذهلة عند اختبار دقته، بالإضافة إلى محرك

أراب فيستا (arabvista) الذي تحول اسمه إلى البحار (www.albahhar.com). وطورت شركة الحمادي أيضًا محرك بحث العالم العربي (www.awse.com) الذي تعتمد آلية البحث فيه على الدليل الرئيسي أو مربع البحث، حيث يحتوي هذا الدليل على إحدى وخمسين فئة رئيسية تحتوي بدورها على فئات فرعية بثلاثة مستويات، تتضمن جميع المواد والمواضيع، لتقدم بذلك المعلومة الشاملة والدقيقة في زمن قياسي للمستخدم العربي؛ حيث تمكن المعهد النرويجي للتكنولوجيا أخيرًا من الوصول إلى أفضل وأشمل محرك بحث على شبكة الإنترنت من وجهة نظر أغلب الباحثين، فهو يضم أكثر من 200 مليون عنوان URL في قاعدة بياناته. إضافة إلى ذلك ظهرت على الساحة خلال السنوات الأخيرة محركات بحث عربية طورتها شركات عربية ومنها محرك «نسيج» و«أين». ولا شك أن الموقع العربي العملاق «باب» www.bab.com يقدم أكبر دليل عربي توصيفي لمجموعة مواقع يصل عددها أكثر من 10000 موقع عربي وأجنبي، مصنفة بطريقة خاصة بحيث تغطي أغلب المجالات. ويعمل محرك البحث استنادًا إلى جميع الكلمات الموجودة في توصيف المواقع، وهو ما يعطي دقة وخيارًا أشمل للبحث، فضلاً عن إمكانية إيجاد الموقع المطلوب عن طريق اسمه أو عن طريق عنوانه URL، حيث إن هذه الميزة ليست متوفرة في كثيرا من محركات البحث الأخرى.

عادة ما تعتمد محركات البحث على نوع واحد من البرامج المستخدمة التي يتشابه أداؤها في كل أنواع المحركات، وسوف يكون التقسيم هنا لإبراز الأنواع ثم لبيان الجوانب التي يمكن أن تتناول أنواع محركات البحث عن طريقها. والجانب الأول هو التخصص الموضوعي؛ ويعني به مدى تخصص محركات البحث وتوقفها على بحث مواقع الموضوع الواحد أو البحث بصفة عامة. أما الجانب الثاني لتقسيم محركات هو البحث في محركات البحث ذات الأشكال وأوعية المعلومات المحددة مثل البحث عن الوسائط المتعددة أو التخصص في نوع واحد من المعلومات سواء

المسموعة أو المرئية. وترتكز هذه الدراسة على تحليل أنواع التقسيم الثاني لمحركات البحث:

1- **من حيث التخصص الموضوعي:**
- محركات بحث متخصصة.
- محركات بحث عامة.
- قواعد البيانات على العنكبوتية.
- محركات بحث متعددة.

2- **من حيث التخصص اللغوي:**
- محركات بحث محددة اللغة.
- محركات بحث متعددة اللغات.

3- **من حيث التخصص الجغرافي:**
- محركات بحث المناطق.
- محركات بحث إقليمية.
- محركات بحث شاملة.

4- **من حيث أساليب الاسترجاع:**
- محركات بحث المواقع.
- محركات بحث المحركات.
- محركات بحث قواعد البيانات.

5- **من حيث التخصص النوعي "وسيط المعلومات":**
- محركات بحث عامة.
- محركات بحث الفيديو.
- محركات بحث المواد المسموعة.
- محركات بحث الصور.

رابعا: محركات بحث الوسائط المتعددة

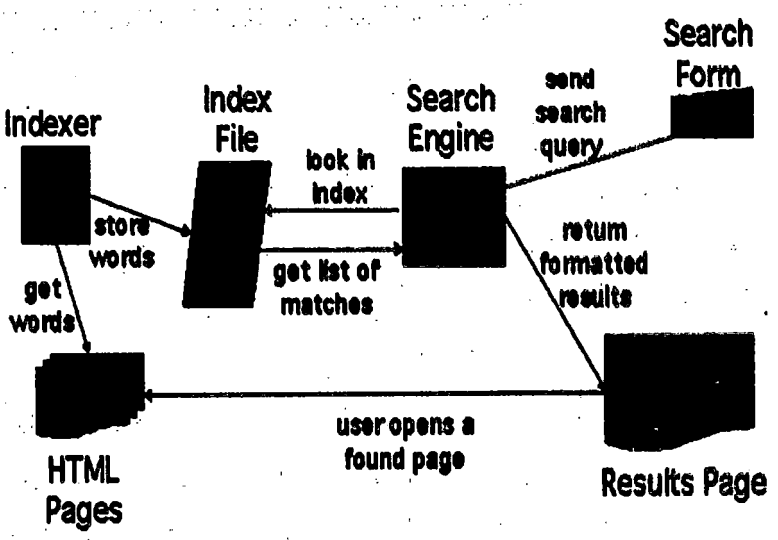
الوسائط المتعددة ترجمة لكلمة لاتينية تتكون من جزئيين "Multi" وتعني التعدد "media" وتعني الوسيطة التي نتناول بها المعلومات (Multimedia). جاءت الوسائط المتعددة بداية لغرض الجمع بين معدات متعددة توظف قدراتها وتحقق الاتصال بينها عن طريق رابط مركزي محوري تعمل من خلاله أو مستقلة عنه (بنظامه) أو يتصل بعضها ببعض عن طريقه، وكان هذا الرابط المحوري هو جهاز الحاسوب الذي يقوم بوظائفه المعتادة وزيادة هذه الوظائف بما يستجد من تطورات. ولا شك أن تعامل الحاسوب مع الأرقام حسابا ومع الكلمات تنسيقا ومع البيانات الرقمية والحرفية فرزا وتبويبا وتصنيفا يختلف عن تعامله مع صور المرئيات عرضا ومطابقة وتحريكا وتسجيلا، ومع الأصوات حفظا وسماعا وإصدارا، يزيد من اختلاف وتنوع عملية البحث على الوسائط المتعددة[6].

أولا: مراحل عمل محركات بحث الوسائط

يعتمد عمل محركات البحث العامة ومحركات بحث الوسائط الخاصة على ثلاث مراحل للتعامل مع صفحات العنكبوتية من جانب والباحثين من جانب آخر. وكل مرحلة من عمل محركات البحث تستخدم ثلاثة برامج هي الزاحف أو العنكبوت في المرحلة الأولى Spider، المكشف في المرحلة الثانية Indexer، وآليات البحث في المرحلة الثالثة Search tools. وتعمل المراحل الثلاث على النحو التالي:

تقوم المرحلة الأولى بنظام الاسترجاع أو محركات البحث تتمثل في تجميع وإضافة مصادر المعلومات التي يقدمها للباحثين بعد تنظيمها، ويتولى برنامج الزاحف Crawler إضافة مصادر المعلومات إلى قاعدة بيانات محرك البحث، ويتبع برنامج الزاحف سياسة إضافة مواقع العنكبوتية إلى قاعدة بيانات محرك البحث التي تهتم في أبسط صورها بنوع موقع المعلومات التي يقف برنامج الزاحف على موضوعات متخصصة وتجميع ما

يصل إليه. ولطبيعة محرك البحث تأثير على عمل الزاحف من حيث اقتناء معلومات العنكبوتية وتشعبها؛ فمحرك بحث الصور - على سبيل المثال - ينعكس على عمل الزاحف في تجميع ملفات الصور الرقمية فقط دون غيرها من الملفات الأخرى. ويستخدم برنامج الزاحف طبيعة تكوين وترابط نسيج العنكبوتية الرابط بين مواقعها وصفحاتها؛ حيث ما يلبث برنامج الزاحف أن يصل إلى المواقع على الشبكة العنكبوتية حتى يخلل كل الروابط Links التي توجد في هذا الموقع ليصل منها إلى غيرها من المواقع والصفحات، ويقوم برنامج الزاحف بوضع نسخة من هذا الموقع في قاعدة البيانات ليؤدي بها برنامج المكشف عمله في المرحلة التالية للإضافة. ولا يقف برنامج الزاحف عند زيارة مواقع العنكبوتية مرة واحدة وإنما يضع البرنامج دوزة كل فترة محددة يرجع خلالها إلى نفس المواقع والصفحات التي التقطها من قبل لمراقبة التحديث الذي يمكن أن يكون قد شمل هذه الصفحات، وعادة ما تكون فترات التحديث قريبة في اغلب محركات البحث[7].



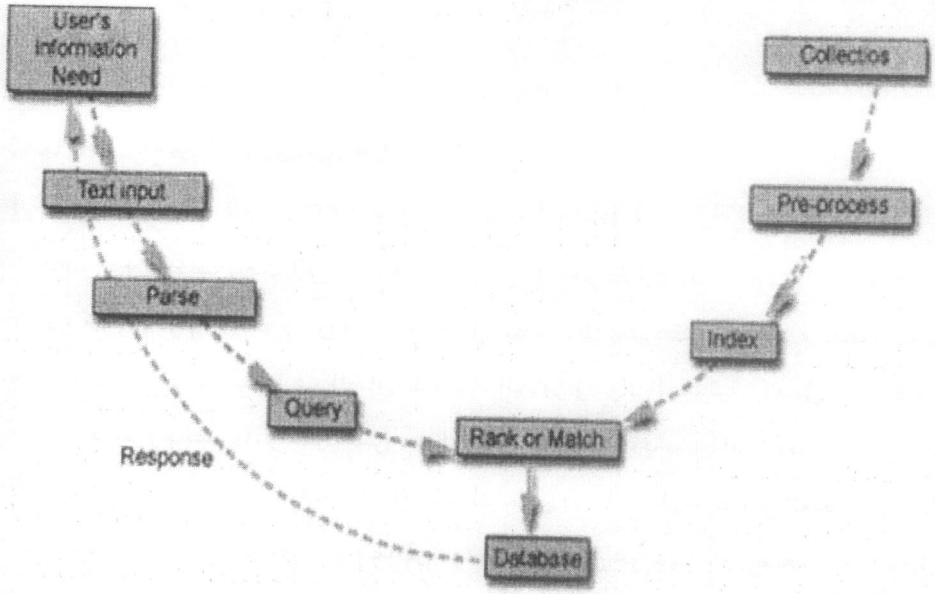
شكل رقم (1) يوضح عمل برنامج المكشف في دائرة عمل محرك البحث.

كما يوضح الشكل رقم (1)، تنتهي المرحلة الأولى بإضافة نسخة من الموقع إلى قاعدة بيانات محرك البحث، ليبدأ برنامج المكشف في أداء مهامه في تكشيف هذه المواقع وإخراج مجموعة الكلمات الكشفية التي تساهم في استرجاع هذه المواقع والصفحات بحسب الحاجة الموضوعية لها في مرحلة البحث. والبرنامج الثاني في محركات البحث الذي يعرف بالمكشف أو المفهرس Indexer or Cataloger، يكمن عمله في إعداد تسجيلات ببليوجرافية لمواقع العنكبوتية حديثة الإضافة إلى قاعدة بيانات أداة البحث، وتخزن هذه الإضافة وفقا لما ورد في جزء المكتبات والمعلومات عن (التمثيل المكثف)؛ حيث لا يتم تخزين موقع العنكبوتية بأكمله في المحركات، وإنما يتم ربط الحقول والبيانات بمسار هذا الموقع لسهولة الوصول إليه، وكل ذلك يتم في شكل مسجل كما هو الحال في الفهارس الآلية للمكتبات OPAC. إن عملية الفهرسة قد تتم بواسطة الإنسان أو برامجة آلية وذلك حسب نوع أداة البحث، أما في حالة التكشيف والفهرسة الآلية فان برامج التكشيف الآلي تستخدم بنية صفحة العنكبوتية web page structure في التعرف عليها ووضع كلمات أو مصطلحات معبرة عن الموضوع أو الموضوعات التي تتضمنها الصفحة العنكبوتية، ومن خلال البرنامج يتم فحص كل أجزاء الصفحة التي تحتوي على كلمات حول الموضوع الموجودة بالعنوان على شكل كلمات نصية أو مفتاحية التي تم وضعها من قبل المصمم لصفحة العنكبوتية في رموز (meta tags)، وتجمع هذه الرموز والكلمات لتدل على الموضوع أو الموضوعات في الصفحة العنكبوتية. وبعد تفحص كل الجوانب السابقة يتم وضع كلمات مفتاحية داخل المسجل الموقع أو وضع كل كلمات النص الطبيعية في وضع نشط ليتم مقارنتها بعد ذلك مع كلمات البحث التي يضعها الباحث عند إجراء البحث.

إن العديد من قواعد الوصف الببليوجرافي لمصادر العنكبوتية بدأت في الظهور، وقد أقرت المنظمة القومية لمعايير المعلومات The National Information

ومعهد المعايير الوطنية Standards Organization (NISO) and الأمريكى American National Standards Institute (ANSI) معيار دبلن المحورى Dublin Core، والمتعلق بفهرسة الصفحات العنكبوتية وبناء data meta الخاصة بها، وهو مكون من خمسة عشر عنصرًا: هي: Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier (URL), Source, Language, Relation, Coverage and Rights (copyright information). وقد صدر هذا المعيار تحت رقم Z39.85 ، ويمكن لمحركات البحث أن تعمل وفق هذا المعيار في التكشيف والبحث للأشكال المختلفة من الوسائط المتعددة والذي يشكل في حد ذاته مجالاً خصبًا للدراسة في الوقت الحاضر[8].

يختلف التعامل في المرحلة الثالثة عن المرحلتين السابقتين؛ حيث يتعامل محرك البحث مع المستفيدين وليس مع صفحات العنكبوتية. ترتكز هذه المرحلة على إعداد مجموعة من استراتيجيات البحث مكونة من مجموعة من آليات البحث ليتعامل بها الباحثون عن المعلومات وقاعدة بيانات والمحرك لاستدعاء مصادر معلومات العنكبوتية، ويأخذ هذا الجزء من عمل محرك البحث الأهمية الأكبر من الجزأين السابقين؛ نظرا إلى أنه الواجهة التي يتعامل معها المستفيد من كل أجزاء محركات البحث والذي يبني عليه المستفيدون تقييمهم لأداء محركات البحث، بالقدر الذي تكون عليه آليات البحث من كفاءة يكون أداء المحرك في تحقيق التحقيق Precision المطلوب والاستدعاء المطلوب، الذي يختلف من مستخدم لآخر.



شكل رقم (2) رسم يوضح الإجراءات الكاملة لمحركات البحث.

الشكل رقم (2) يوضح المهام التي يقوم بها محرك البحث بالتعامل بين المستخدم وقاعدة البيانات؛ حيث يوضح الإجراءات التي تتم قبل إضافة مواقع العنكبوتية وصفحاتها، والإجراءات التي تتم من قبل المستخدم عند الحصول على معلومات، بداية من الحاجة للمعلومات إلى التعامل مع قاعدة البيانات وطرح الاستفسارات الموضوعية ثم تلقي الإجابات عليها من قاعدة البيانات في صورة مواقع وصفحات العنكبوتية المضاهية للموضوع المطلوب. إلا أن الباحث يرى أن التسميات التي جاءت في هذا الشكل غير مناسبة بعض الشيء؛ فعلى سبيل المثال مصطلح INDEX يمكن أن يطلق على قاعدة البيانات نفسها، فضلا عن استخدام مصطلح سابق لهذا المصطلح وهو Pre-Process أي العمليات التي تتبع التجميع وتسبق الإضافة إلى قاعدة البيانات التي تتم فيها عملية التكشيف. وعلى الرغم من ذلك فإن هذا

الشكل هو المعمول به في إجراءات محركات البحث عند ضم مواقع العنكبوتية وصفحاتها.

ثانيا: أنـــواع محركات بحث الوسائط

1. محركات بحث الفيديو Video Search Engines

تمثل محركات بحث الفيديو أحد أهم أنواع محركات بحث الوسائط المتعددة، إلا أنه لا يوجد لهذا النوع من وسائط حمل المعلومات محركات بحث متخصصة في بحث ملفات الفيديو كما هو الحال في بحث الصور والمواد الصوتية. وتعمل محركات بحث الفيديو بنفس المراحل التي تعمل بها محركات بحث النصوص والوسائط الأخرى؛ إذ يقوم محرك البحث بالاعتماد على برنامج الزاحف أو العنكبوت في استقطاب ملفات العنكبوتية والتقاطها والتعرف على أنواعها ثم وضعها في المكان المخصص لها، من خلال الاعتماد على برنامج المكشف في تحليل النص التابع لمادة الفيديو؛ من مسار URL وعنوان وكلمات واصفة Meta Tags، ليخرج منها برنامج الزاحف بمجموعة الكلمات التي يستخدمها الباحثون في استدعاء ملف الفيديو. هذا من ناحية أداء المحرك نفسه، إلا أن الأمر الآخر هو ما يخص المحتوى نفسه فإن محاولات عديدة تعمل لإيجاد معايير لتحليل ملفات الفيديو واسترجاعها؛ سواء الوصف المادي أو الموضوعي.

وقد عمل معيار Dublin Core على التقعيد للوصف المادي والموضوعي لتشعب معلومات الفيديو، كما عمل في السابق على الصور، حيث يأخذ الوصف هنا اتجاه التقسيم لتحليل موضوع المادة إلى أقسام أو مقاطع ثم تقسيمها إلى مشاهد والمشاهد تقسم بدورها إلى لقطات ثابتة، يأتي ذلك في دراسة Jane Hunter[6] عن إقامة نظام للكشف والتصفح واسترجاع يعمل في مكتبات المواد السمعية البصرية، مراعيا أن يكون الوصف المادي لمواد الفيديو الرقمية شاملا كل قسم من الأقسام المكونة لمادة

الفيديو. وقد كان الوصف الكامل لمحتوى الفيديو الرقمي تبعا لمعيار Dublin Core كما يلي:

جدول يوضح عناصر الوصف المادي لمواد الفيديو الرقمية على الإنترنت

النص	البيانات " الحقول " Meta data
عنوان المادة الفيلمية	Title =
القائم على إنتاج المادة الفيلمية	Creator = Producer
موضوع المادة في كلمات محدودة	Subject =
وصف الموضوع في عبارات شارحة	Description =
الناشر	Publisher =
1. تاريخ النشر	Date =
2. نوع المادة الفيلمية	Type = "Image. Moving. Film. Documentary"
وصف الشكل المادي "مقابل حقل التوريق"	Format = 1 videocassette (27 min.) : sd., col. ; 1/2 in
3. محدد الحصول على المادة	Identifier =
4. مصدر المادة الرقمية	Source =
5. اللغة	Language =
6. الأجزاء المكونة للمادة الفيلمية الرقمية	Relation.HasPart = scene1, scene2, scene3, scene4, scene5,...
7. التغطية للمادة " الزمنية والمكانية "	Coverage =

إلا أن التحليل الموضوعي لمواد الفيديو مازال يقف في منطقة ضبابية، تحتاج إلى الكثير من الجهد للكشف عن الطرق الفعالة في تكشيف وتحليل موضوع مثل هذا النوع من الوسائط المتعددة.

2. محركات بحث الصور Search Engines Image

تتفوق محركات بحث الصور على محركات بحث الوسائط الأخرى في التواجد العددي على الشبكة العنكبوتية، كما أن محركات بحث الفيديو يعتريها التخصص في التعامل مع ملفات الصور سواء تلك العامة "الطبيعة" أو الإعلامية أو الصور العلمية. وتتميز أدوات العنكبوتية المتخصصة في بحث الصور عن الوسائط الأخرى بأنها تتكون من فئتين متباينتين في العمل هما: محركات بحث الصور، وقواعد بيانات الصور. إن كلا النوعين يختلف عن الآخر في طبيعة التعامل مع الصور من ناحية التحليل المادي والموضوعي، إضافة إلى أن طبيعة التعامل في كل نوع تختلف عن الآخر بطبيعة مفهوم محرك البحث الذي يعمل بطريقة آلية خالصة تعتمد على برامج ثلاث في إضافة وتنظيم وبحث ملفات العنكبوتية على اختلاف أنواعها الموضوعية والشكلية، خلاف ما توجد عليه قواعد البيانات من التدخل البشري في معظم إجراءاتها؛ مثل تحديث قاعدة البيانات وإضافة المعلومات إليها، وتنظيم مجموعات الملفات التي توجد فيها، خلاف مرحلة البحث فإنها، بالإضافة إلى أن محركات البحث تعتمد على استراتيجيات بحث، تتكون من آليات مختلفة.

تتساوى الصور الرقمية على العنكبوتية مع أنواع الوسائط الأخرى في الحاجة إلى التوصيف المادي والموضوعي، ومن الممكن أن تكون الصور الرقمية هنا مثالا لتناول التحليل المادي والموضوعي للوسائط الأخرى؛ حيث إن جميع ملفات الصور الرقمية على العنكبوتية، مع اختلاف موضوعاتها "العلمية والطبيعة والإعلامية"، تتساوى في الوصف المادي للخصائص الرقمية وحقول الترميز داخل ملفات لغة الترميز المعيارية HTML، إلا أن التحليل الموضوعي يختلف من نوع إلى آخر من الصور. كما هو بالمثل على مستوى كل نوع من الوسائط؛ الفيديو أو الصوت فان ملفات هذا النوع تتساوى في الخصائص الرقمية، مما يمكن معه إيجاد معايير لوصف هذه الملفات وتحليلها. إلا أن الصعوبة في تناول مختلف الأنواع تأتي في التحليل الموضوعي لها فإن

الصورة التي تعتمد على حاسة النظر تحتوى على ثلاثة مستويات من التحليل الموضوعي؛ الأول يتمثل في العناصر الأساسية المكونة للصورة التي تقع عليها العين أول ما ترى الصورة، والثاني فهو العناصر الثانوية المكملة للصورة وتكون في خلفية الصورة، أما المستوى الثالث فهو الخاص بالمعنى والمفاهيم التي تحويها الصورة والتي تكون كل العناصر في التعبير عنها. وبذلك فإن الصوت والفيديو ينتظران من يتبنى البحث بهما في مواصفات الشكل والمحتوى.

3. محركات بحث الصوت MP3 Search Engines

تعد ملفات الصوت MP3 هي أشهر وسيط للتعامل مع الموسيقى على الحاسوب، وبالإضافة إلى ترميز الصوت فإن هذه الملفات تتسم أيضا بوجود مدى محدود من البيانات، هذه الملفات تسمح من خلاله بإضافة نص أو كلمات وحقول مصاحبة لملفات الصوت Meta data، وتوجد لمثل هذه الملفات العديد من الأشكال Formats الّتي تحتوي على حقول لوصف المادة الصوتية باستخدام الكلمات المفتاحية؛ ومنها:

ID3v1

إن ملفات الـMP3 في شكلها الأصلي لا تحتوي على مساحة لكتابة حقول الوصف، إلا أن الباحث Eric Kemp قد حرص على إرفاق مثل هذه الحقول بالمادة الصوتية في مساحة بسيطة ليخرج شكل ID3. وتجمع هذه الحقول مختلف البيانات الببليوجرافية عن المادة الصوتية مثل عنوان الأغنية أوالمادة الصوتية، اسم المطرب أو من يحل محله، الألبوم، السنة "تاريخ النشر"، التعليق، ومساحة 1 كيلوبايت لكتابة نوع المادة؛ موسيقى كلاسيك، روك Rock، رقص، وطنية... الخ. وما لبث أن تطور المعيار على يد Michael Mutschler إلى شكله ID3v1 بإضافة معلومات عن رقم مسار الصوت على القرص المليزر. إلا أن الباحث قد لاحظ في هذه الحقول غياب الحقل أو الترميز الخاص بالموضوع ووصف محتوى المادة المسموعة.

ID3v2

قامت مجموعة من الأشخاص على رأسهم Martin Nilsson في عام 1998، بوضع مجموعة من الحقول في أشكال تناسب كل حقل على حدة، ومن خلالها توضع البيانات الببليوجرافية الخاصة بالمادة المسموعة وهي ID3v2، إلا أنها توضع قبل الصوت في الملف الحاوي لها بعكس ما كان عليه [7]ID3v1.

4. محركات بحث المواد الحركية Animation Search Engines

ارتبط ظهور المواد الحيوية باستخدام الحواسيب في التعليم والبحث؛ حيث لا يوجد لمثل هذه المواد أصول تقليدية مثل باقي الوسائط الأخرى. ولا تلقى مواد الرسوم الحيوية Animation نفس المقدار من الاهتمام الذي تناله الأنواع الأخرى من محركات بحث العنكبوتية، فلا تتيح محركات البحث إمكانية البحث المتخصص في ملفات الرسوم الحيوية، إلا قليل القليل الذي يبحث داخل هذا النوع من الوسائط. ويرجع الباحث ذلك إلى قلة تواجد هذا النوع من الملفات على العنكبوتية؛ ذلك أنّ محركات البحث لا تعمل مع مختلف أنواع الملفات التي توجد على العنكبوتية بنفس المعايير والإجراءات، مما يعني أن قصور بحث هذا النوع يرجع إلى تواجد ملفات الرسوم المتحركة في الشكل الرقمي في نطاق بيئة العنكبوتية.

نمـــاذج من محركات بحث الوسائط المتعددة

أولا: محركات بحث وسائط عامة:- General Media Search Engines

Lycos

يشتمل محرك "www.lycos.com" على مجموعة كبيرة من ملفات الوسائط المتعددة، إلا أنه لا يعتمد على آليات بحث متقدمة كتلك التي توجد في غيره من محركات البحث الأخرى، يسير هذا المحرك وفق طبيعة عمل محركات البحث العامة بالاعتماد على عناصر عمله الثلاث، ألا وهي برنامج الزاحف Crawler أو سفير التعارف لمواقع العنكبوتية المختلفة وصفحاتها، ثم التحليل الأولى لأنواع تلك الصفحات

التي تصنفها إلى صفحات نصية تخضع للبحث العام أو تلك التي تندرج أسفل أحد أنواع ملفات الوسائط المتعددة. ثم تكشف تلك الصفحات بالتعرف على أهم الكلمات لدالة على محتوى المادة سواء المرئية أم المسموعة، يلي ذلك دور آليات البحث في استدعاء تلك الملفات بحسب الكلمات المفتاحية الداخلة في استراتيجية البحث.

AltaVista

يعتمد محرك AltaVista كما سبق على آليات بحث متقدمة تختلف حسب طبيعة نوع وسط المعلومات موضع البحث. وعلى سبيل المثال فإنه في حالة بحث ملفات الصوت يقدم آليات اختيار نوع الملف AIFF أو AU أو WAVE، إلى جانب مدة تشغيل ملف الصوت. أما في حالة بحث الصور Images فإنه يقدم آليات مثل لون الصورة، نوع الصورة مثل الصور الفوتوغرافية أو الرسومات.

Fast

تتبع أداة البحث "http://www.multimedia.alltheweb.co" Fast محرك البحث المتعدد all the web؛ حيث يعد Fast الوسيط المسؤول عن بحث ملفات الوسائط المتعددة من صور وفيديووصوت. ولا تختلف هذه الأداة في بحث ملفات لوسائط عن مثيلاتها السابقة من محركات البحث في العمل سواء أكان التجميع لهذا النوع من الملفات أم كان التكشيف، إلا أن الاختلاف قد يرجع فقط إلى تغير بعض آليات البحث، وهو الأمر الذي تتسم به محركات البحث عامة.

ثانيا: محركات تدعم نوع واحد Specialized Media Search Engines

Ditto

يعد "www.ditto.com" Ditto من أهم محركات بحث الصور على الإنترنت، وهو من المحركات المتخصصة في شكل الصور. ويبحث المحرك في أنواع من الصور مثل الصور الفوتوغرافية والرسومات واللوحات الفنية وغيرها من العناصر المكملة لمصطلح Image. ويستخدم Ditto برنامج الزاحف في تكوين مجموعاته والإضافة إلى

قاعدة البيانات، حيث يحدد برنامج الزاحف محتوى الصفحات حديثة الالتقاط وهو الإجراء الذي يعتبر أول مراحل تحليل محتوى صفحات الصور، والتعرف على مجموعة الكلمات التي يمكن استرجاع الصور بها. هذا إلى جانب الاعتماد على تحديد مدى الصلة لمجموعة الصور موضع الاستدعاء، مما يكون من شانه تحقيق درجة عالية من الدقة في بحث المستفيد. إن ما سبق يعني استخدام المحرك لمجموعة الكلمات التابعة للصورة في العرف على محتوى الصورة الموضوعي. ويتيح محرك البحث التصفح الموضوعي العام عن الاسترجاع، إضافة إلى إمكانية البحث بالكلمات المفتاحية، وإن كان الباحث يعتقد أنها أسلوب الاسترجاع الرئيسي لصور قاعدة البيانات[8]. أما عرض النتائج فإنه يعتمد على شكلين أساسيين هما عرض الصور بجانب البيانات التابعة لها، والآخر هو عرض وحدات الصور التي تضاهي كلمات البحث كلقطات thumbnails يمكن للتعرف على الصور ومن ثم يمكن الانتقال إلى الشكل الأصلي للصورة.

Find Sounds

يعتبر أحد محركات البحث التي تختص بالبحث عن المواد الصوتية، يحاكي محرك www.findsounds.com مختلف الاستخدامات لكل أعمار مستخدمي العنكبوتية. ويعتمد محرك البحث على برنامج الزاحف كطبيعة محركات العنكبوتية في اقتناء صفحات العنكبوتية، إلا أنه يعمل بشكل مخصص على ملفات المواد الصوتية، فضلا عن اختيار أنواع محددة من الملفات الصوتية مثل AIFF, AU, and WAVE التي تستخدم مع أرضيات تشغيل مثل Macintosh, Unix, and Windows، والتي يشيع استخدامها على العنكبوتية. ويعتمد هذا المحرك أيضا على الكلمات التابعة للمادة الصوتية في ملفات الصوت على العنكبوتية لاستخدامها في عملية التكشيف ومن ثم استدعاء الصوت بحسب كلمات البحث المدخلة من جانب المستخدم. أما عرض النتائج فإنه يتمثل في عرض العشر وحدات

الأولى من مجموع النتائج التي طابقت كلمات البحث، إلى جانب ذلك فإن واجهة النتائج تتضمن خيارات الانتقال إلى صفحات النتائج الأخرى أو فتح ملف الصوت، عرض مسار ملف الصوت على العنكبوتية Web URL إلى جانب حجم المادة الصوتية التي لا تزيد لهذا المحرك عن 2 M كما جاء في خيارات البحث[9].

Corbis

يعد موقع البحث www.corbis.com أهم مواقع بحث الصور على العنكبوتية، ويرجع ذلك إلى عوامل عدة من أهمها حجم المجموعات التي تخضع للبحث على الخط المباشر، وتقدر هذه المجموعة بـ 2.1 مليون صورة من حجم كلي يصل إلى 65 مليون صورة. ويؤدي هذا الموقع عمله في بحث الصور من خلال تجميع كم كبير سواء على الخط المباشر أو قواعد البيانات المكتنزة، وهو يهدف اكثر ما يهدف إلى الربح. ولم يشر الموقع إلى إجراءات العمل إلا أنه لا يعد محرك بحث كما يعكس المفهوم من طبيعة أداء محركات البحث، وقد جاء ذلك بالنظر إلى طريقة تجميع المواد المصورة التي تعتمد في معظمها على الإضافة اليدوية إلى قاعدة البيانات. ومن خلال مجموعة البيانات التي ترافق عرض الصورة في التسجيلة النهائية المكونة من المادة المصورة، جملة واصفة للصورة في شكل عنوان، ثم المصور، ورقم الصورة. إلا أن هذا المحرك الاعتباري يسمح باستخدام أسلوبين من البحث هما البحث بالتقسيم الموضوعي أو التصفح Browsing لرؤوس الموضوعات، ولا يتبع هذا التقسيم أي خطط تصنيف محددة إلا أنه يعبر عن الموضوعات بشكل عام. أما الأسلوب الآخر فهو المتمثل في البحث بالكلمات المفتاحية، مع إمكانية استخدام آليات بحث مثل؛ استخدام علامات التنصيص " " لتطابق المصطلحات كما كانت في شكل البحث، إضافة إلى الاستخدام التلقائي لمعاملات البحث البوليني AND, OR. كل ما سبق يعمل على الوصول إلى درجة عالية من الدقة المطلوبة في البحث.

نتائج الدراسة

تخلص هذه الدراسة يحدد النتائج التالية:

1. إن التعدد والتنوع في أشكال وسائط المعلومات الرقمية المرئية والمسموعة التي أفرزتها بيئة الويب، يحتاج إلى مزيد من الدراسات والبحوث في هذا الجانب وبخاصة محركات البحث باللغة العربية.
2. إن التنامي السريع لمحتوى الإنترنت وتطور وسائطه النصية والصوتية والمرئية وتعددها، يؤكد الحاجة إلى أكثر من منظمة لرعاية تنظيم المعلومات ووضع المعايير اللازمة.
3. تعمل جميع محركات البحث باستخدام آلية واحدة لتنظيم وبحث مصادر الويب، وذلك على الرغم من اختلاف وسائط حمل المعلومات وتعددها داخل الوسيط الواحد. وأكد ذلك التوحد في بناء محركات البحث اعتمادها على ثلاثة برامج فرعية لجمع وتنظيم وبحث معلومات الويب المختلفة، وتعمل هذه البرامج بآليات مخصصة لاسترجاع النصوص.
4. نظرا لأن محركات بحث الصور تعمل بطريقة تفاعلية غير ثابتة (ديناميكية)، فإنه لا يصلح تطبيق أساليب قياس الكفاءة في نظم استرجاع الثابتة على محركات بحث الوسائط المتعددة.

الخاتمة

ما سبق عرضه حول محركات بحث الوسائط المتعددة، ليس سوى خطوة أولى على طريق معرفة محركات البحث واستخدامها بصورة مثلى من قبل المتخصصين وغير المتخصصين في هذا المجال، ويعرف بعض أدوات تنظيم المعلومات على الشبكة العنكبوتية، وبالأخص محركات بحث الوسائط المتعددة، كما أنه يمكن أن يطور من العمل على محركات بحث العنكبوتية عامة، التي تختلف بدورها وطبيعتها عن

قواعد البيانات على العنكبوتية، بخلاف الوجه الآخر للبحث على العنكبوتية المتمثل في أدلة البحث Search Directories. وهو ما يحاول الباحث العمل به وتعريفه من خلال هذه الدراسة. وبتناوله أوعية المعلومات ذات الأشكال المختلفة على العنكبوتية مثل الفيديو والصور والصوت، التي مازالت في حاجة إلى التوصيف والمعايرة، سواء في أشكالها التقليدية أو الشكل الرقمي في بيئة الشبكة العنكبوتية.

المراجع كتب ومواقع:

[1] داليا نصار رياض. محركات البحث العربية على الإنترنت. ـ مكتبات ـ نت. ـ مج 2، ع1، 2 (2005).

[2] محمد زايد. محركات البحث العربية فى شبكة الإنترنت. ـ المجلة العربية للمعلومات.. مج20، ع 2(1999).

[3] بهاء شاهين. مرجع سابق. ص 332.

[5] فهد بن ناصر العبود. آلية البحث فى الإنترنت: محركات البحث أنواعها، مهامها، طرق البحث فيها / عرض وتحليل أ.حمد على تمراز .. مجلة مكتبة الملك فهد الوطنية .. مج 7، ع2(2002).

[6] د. محمد المدني. دور الوسائط المتعددة في التعليم والتدريب، مجلة الاتحاد العربي للتدريب التقني. العدد الثالث والرابع (2009).

[7]هشام محمود عزمى. مواقع المكتبات والمعلومات: دراسة تحليلية لشبكة الإنترنت ـ مجلة المكتبات والمعلومات العربية ـ س17، ع4 (1997).

http://www.firstmonday.dk/issues/issue2_2/peterson/

[8] Cosme, Gallianno. How Do Search Engines Work? **ISEdb COM, 2005,** http://www.isedb.com/db/articles/1174/1/How-Do-Search-Engines-Work?

[9] Hunter , Jane and Newmarch, Jan. An Indexing, Browsing, Search and Retrieval System for Audiovisual Libraries , Uni. of Queensland , 2004

[10] Ditto : Privacy Policy , Ditto , 2002 , http://www.ditto.com/privacy.asp

11-Al-Kharashi, Ibrahim A. A Web Search Engine for Indexing, Searching and Publishing Arabic Bibliographic Dat:abases, King Abdul Aziz City for Science and Technology, Computer and Electronics Research Institute, [2001] at:http://www.isoc.org/inet99/proceedings/posters/085/

12-Cohen, Laura. Conducting Research on the Internet, university at: Albany libraries, 2003, visited at: 1/6/2003, : www.uah.edu/502/Elmore_POTW_Ch4_report.

13-Drabenstot, Roren M.. Web search strategy development,gale group, (2001), visited (2/4/2002), : http://www.sciencemag.org/cgi/content/abstract/

14-How Search Engines Work. Search Engine Wat:ch, 2002, http://www.searchenginewat:ch.com/webmasters/article.php/2168031
15-Lazewski, Barbara. WWW Search Strat:egies, University of Wisconsin-Madison, 2006, http://www.library.wisc.edu/libraries/Steenbock/services/wwwstrat:.htm