البحث في جريان المعلومات

مؤنس طيبي

تضمّ شبكة الإنترنت ملايين المواقع التي تحتوي على كمّيّات هائلة من المعلومات ، تُؤخذ من قَواعد بيانات مختلفة . هذه المعلومات تتضمّن مقالات علميّة ، برامج ، أفلام وصور ، كتب ، صحف ، مجموعات نقاش ، مكتبات ، متاحف ، مراكز تعليميّة ، مراكز شراء وغيرها من الكثير . هذه القائمة تمثّل جزءًا صغيرًا ممّا يتمّ نشره عبر هذه الشبكة العالميّة . من خلال هذه الحقيقة يتّضح لنا أنّه من الصعب جدّا العثور على معلومات معينة بواسطة التنقّل (البسيط) من موقع إلى آخر والبحث عن هذه المعلومات المطلوبة . لذلك هناك حاجة إلى أدوات تساعدنا في البحث عن معلومات . أحد هذه الأدوات يطلق عليها اسم محرّك البحث (Search engine) . محرّك البحث عبارة عن حاسوب خادم أو أكثر ، يُفَهْرِس محتويات وعناوين مواقع شبكة الإنترنت (WWW) في قواعد بيانات ويصدر في صفحة نتائج البحث عناوين صفحات المواقع التي تتطابق مع موضوع البحث المطروح من قبل المستخدم لمحرّك البحث . يحتوي محرّك البحث على برامج خاصّة ، تقوم بتفتيش مواقع شبكة الإنترنت بشكل مستمرّ وتزويد قواعد البيانات بالمعلومات عن هذه المواقع .

خدمات للبحث عن معلومات :

عند البحث عن معلومات في شبكة الإنترنت نستعين بخدمات مختلفة للبحث . هدفها الأساسيّ هو غربلة مواقع شبكة الإنترنت لإيجاد المواقع المطابقة لموضوع البحث . هذه

الخدمات يمكن تقسيمها إلى الأنواع التالية .

- ❐ الدلائل .
- ❐ محرّكات البحث .
- ❐ محرّكات البحث المتحوّلة أو المتوازية .

هذه الأنواع من الخدمات تعمل بطرق وأساليب مختلفة . يجب معرفتها من أجل تحقيق نجاح ونجاعة أكبر لعمليّة البحث .

الدلائل (Catalogs) :

هذه المواقع تحتوي على دلائل مقسّمة إلى فئات مواضيع ، تُجَمَّعُ وتُصَنَّفُ فيها عناوين المواقع الموجودة في شبكة الإنترنت بشكل يدويّ من مجموعة أشخاص . هذه المجموعة من الأشخاص تمثّل هيئة التحرير للموقع وهي المسؤولة عن عمليّة الجمع والفهرسة لعناوين المواقع ، حيث تبحث في شبكة الإنترنت عن المواقع الجديدة أو تلك التي تمّ تغييرها وتصنّفها داخل الفئات المناسبة في الدليل ، الذي يتمّ عرضه للمستخدم[4] . عند اكتشاف موقع جديد يتمّ تقييم هذا الموقع وفق محتوياته أوّلاً قبل إدخاله إلى الدليل . في هذه الطريقة من العمل يمثّل العامل الإنسانيّ الحسنات والسيّئات لهذه الطريقة ، إذ أنّ عمليّة التقييم والفرز عمليّة واضحة لهيئة التحرير ، لكن اتّخاذ القرار بالنسبة لتعريف فئة موضوع أساسيّة أو فرعيّة جديدة ليس دومًا سهلاً وواضحًا ؛ فالدليل مقسّم إلى عدّة فئات رئيسيّة ، وكلّ فئة مقسّمة إلى فئات فرعيّة عديدة . وهكذا تستمرّ عمليّة التفرّع حتّى الوصول إلى فئة ، تحتوي على مجموعة عناوين لمواقع ، تبحث في موضوع هذه الفئة المحدّدة . من هنا يتّضح أنّ الدلائل تحتوي على ترتيب هرميّ للمواضيع وأنّ لتعريف وترتيب فئات المواضيع داخل هذه الدلائل أهمّيّة كبيرة

في إيجاد المعلومات المطلوبة .

مواقع الدلائل تلائم أكثر للبحث عن مواضيع عامّة أو عندما يكون موضوع البحث غير واضح ومحدّد من البداية . فمثلاً إذا أردت البحث عن جميع الجامعات في العالم أو في دولة معيّنة عليك النقر على الفئة الرئيسيّة الملائمة لهذا الموضوع ومن ثمّ الاستمرار في عمليّة تحديد وتضييق مجال البحث بواسطة اختيار الفئات الفرعيّة الملائمة لهذا الموضوع . عليك الاستمرار في هذه العمليّة حتى تصل إلى مجموعة عناوين مواقع الجامعات التي تبحث عنها . مثال جيّد لهذا النوع من خدمات البحث هو الدليل المشهور Yahoo وعنوانه (http://www.yahoo.com) .

على سبيل المثال انطلاقًا من موقع Yahoo يمكنك الوصول إلى عناوين الجامعات[1] الموجودة في لبنان بواسطة اختيار الفئات الملائمة بالتسلسل التالي[2]:

- Education
  - College and University
    - Lebanon
      - في هذه الفئة الفرعيّة ستجد عناوين مواقع الجامعات في لبنان

تحتوي مواقع الدلائل إضافةً إلى الدليل نفسه أيضًا على مربّع نصّ ، يسمح للمستخدم بإدخال الموضوع الذي يبحث عنه ، حيث يعرض الدليل فيما بعد عناوين المواقع التي تتطابق مع هذا الموضوع . عمليّة البحث تتمّ إمّا في جميع الفئات التابعة للدليل أو في الفئة المحدّدة من قبل المستخدم وجميع الفئات الفرعيّة التابعة لها .

على المستخدم وقت البحث أنْ يحاول دومًا تضييق وحصر مجال البحث من أجل الحصول على المواقع المنشودة للموضوع الذي يبحث عنه ولكي يستطيع فحص هذه المواقع فيما بعد . عمليّة تضييق مجال البحث في مواقع الدلائل تتمّ بشكل أساسيّ عن

طريق الاختيار الصحيح للفئات الفرعيّة الملائمة .

ممّا ذكر أعلاه نرى أنّ لهيئة التحرير دورًا كبيرًا ومهمًا في تقييم المواقع وأيضًا في دقّة نتائج البحث وذلك لأنّ محتويات المواقع هي القاعدة الأساسيّة لعمليّة الفهرسة المتّبعة في هذا النوع من خدمات البحث وبذلك أيضًا لعمليّة البحث .

محرّكات البحث (Search Engines) :

في محرّكات البحث تتمّ عمليّة البحث والفهرسة لمعلومات وعناوين المواقع بشكل أوتوماتيكي ، حيث تؤدّي برامج خاصّة هذه المهام . يتكوّن محرّك البحث من ثلاثة أقسام أساسيّة [4] :

❐ جامع المعلومات (العنكبوت) :

عبارة عن برنامج[3] يتجوّل من موقع إلى آخر لفحص الشبكة بكلّ أبعادها . وظيفة هذا البرنامج هي جمع المعلومات المتواجدة في المواقع التي يزورها ومن ثمّ إرسالها لمحرّك البحث للفهرسة . تختلف محرّكات البحث فيما بينها أيضًا بالنسبة لكميّة المعلومات التي تنقلها برامج العنكبوت التابعة لها ؛ فهناك برامج تنسخ أجزاء معيّنة من المعلومات الموجودة في الموقع وأخرى تنسخ جميع معلومات الموقع وتصدرها إلى محرّك البحث لعمليّة الفهرسة . تستخدم برامج العنكبوت الوصلات التشعّبيّة (Links) الموجودة في المواقع من أجل الانتقال إلى مواقع أخرى [Haskin, 1998] .

❐ برامج الفهرسة :

بعد وصول المعلومات من برنامج العنكبوت إلى محرّك البحث تبدأ برامج الفهرسة بوظيفتها وهي ترتيب المعلومات وعناوين المواقع في الفهرس ، لكي تصبح قابلة للبحث .

❐ **برامج تقييم الاستعلام :**

هذه البرامج تبعث كلمات البحث أو الاستعلام (Query) إلى الحاسوب الخادم في محرّك البحث ، الذي يحتوي على قواعد البيانات ، ليقوم بدوره بعرض نتائج البحث للمستخدم .

تحتوي جميع محرّكات البحث على مربّع نص يسمح للمستخدم كتابة الموضوع الذي يبحث عنه وذلك من خلال الكلمات التي يراها ممثّلة لموضوع البحث . بعد ذلك يبدأ محرّك البحث بتفتيش قواعد البيانات (المُفَهرَسَة) التابعة له لعرض عناوين المواقع التي تتطابق مع هذا الموضوع . محرّكات البحث تستخدم طرقًا مختلفة لتحديد التطابق بين المواقع والاستعلامات المطروحة من قبل المستخدم . الطريقة الشائعة هي إحصاء عدد المرّات التي ترد فيها كلمة البحث في الموقع . المنطق من وراء ذلك هو الاعتقاد بأنّه كلّما ازداد ظهور كلمة البحث في الصفحة ، كلّما زادت نسبة ترابط الصفحة بالموضوع ولذا يتمّ تصنيف هذه المواقع المحتوية على العدد الأكبر من كلمات البحث في صفحة نتائج البحث كالمواقع الأكثر مطابقة لموضوع البحث [Haskin, 1998] .

هناك طرق أخرى لتحديد التطابق بين موضوع البحث والمواقع ، منها تلك التي تحصي عدد مرّات ظهور كلمة البحث في الأسطر الأولى فقط للصفحة وذلك بسبب الاعتقاد ، بأنّ احتمال تطابق الصفحة مع موضوع البحث هو أكبر عندما تظهر كلمة البحث في مطلع الصفحة . عند برمجة صفحات الإنترنت تستعمل وسوم خاصّة (Meta-tags)، تهدف إلى وصف محتويات الصفحة التي يتمّ قراءتها من قبل عنكبوت أو روبوت محرك البحث . الاختلاف في كميّة ونوعيّة المعلومات المنقولة من المواقع بواسطة برامج العنكبوت إلى محرّك البحث هي أحد الأسباب التي تؤدّي إلى الاختلاف في نتائج

البحث عن نفس الموضوع في محرّكات بحث مختلفة .

أكبر وأشهر محرّكات البحث العالميّة هي :

- http://www.altavista.com
- http://www.hotbot.com
- http://www.northernligth.com
- http://www.webcrawler.com
- http://www.excite.com

للحصول على قائمة جميع محرّكات البحث الموجودة في شبكة الإنترنت بإمكانك زيارة الموقع التالي [Sander-Beuermann, 1998] :

http://www.yahoo.com/computers_and_Internet/Internet/world_wide_web/searching_the_web/

محرّكات البحث المتحوّلة أو المتوازية (Meta-Search Engines) :

هذه المحرّكات تعمل مع مجموعة محرّكات بحث (عاديّة) أخرى ، حيث تبعث بالاستعلام الذي يتمّ إدخاله من قبل المستخدم إلى هذه المجموعة من محرّكات البحث والتي تقوم بدورها وبشكل مستقل ومتوازي بتنفيذ عمليّة البحث . بعد ذلك يبعث كلّ محرّك بحث نتائج بحوثه إلى محرّك البحث المتحوّل والذي يقوم بتقييم ودمج النتائج المرسلة إليه من قبل محرّكات البحث المختلفة ليظهرها للمستخدم . في صفحة نتائج البحث تعرض هذه المحرّكات المتحوّلة عدد المواقع الذي تمّ إيجادها من قبل كلّ محرّك بحث مشترك في عمليّة البحث وبجانب كلّ عنوان تظهر فقرة تصف محتويات الموقع واسم محرّك البحث الذي عثر عليه .

هذا النوع من الخدمات يأتي كحلّ لتقليص الفراغ الناتج عن طريقة عمل محرّكات

البحث العاديّة . ولكن ، يجدر بالذكر بأنّ البحوث العلميّة في هذا المجال تتّجه في الآونة الأخيرة إلى تصميم وتطبيق «نظام الوكيل الذكيّ» (Intelligent Agent System) الذي يساعد مستخدم الشبكة في العثور على المعلومات المنشودة بشكل أنجع وأدق، وذلك عن طريق إعطائه معلومات تصف أهداف ورغبات المستخدم وأيضًا طريقة عمله [طيبي ، 1999] .

أمثلة لمحرّكات بحث متحوّلة [Sander-Beuermann, 1998] :

1. http://www.highway61.com
2. http://www.metacrawler.com

تزايد عدد محرّكات البحث في شبكة الإنترنت يؤدّي إلى تزايد التنافس فيما بينها ، حيث يحاول كلّ محرّك بحث كسب ثقة الزوّار من خلال الخدمات التي يقدّمها ، لكن مع وجود أكثر من ألف محرّك بحث في شبكة الإنترنت ، يظهر طبيعيًا السؤال : **أيّ محرّك بحث هو الأفضل** ؟ لا أعتقد أنّه يوجد إجابة واضحة وصريحة على هذا السؤال ، حيث يجب طرحه بشكل آخر وأدقّ ، كالتالي : **أيّ محرّك بحث هو الأفضل من أجل إيجاد المعلومة** X ؟ لمعرفة ذلك علينا أنْ نعرف بعض الأشياء الأخرى عن محرّك البحث ، مثل [Harris, 1997] :

- ما هي كميّة المعلومات المخزّنة في محرّك البحث ؟
- ما هي المعلومات التي تأخذ من المواقع ، لكي تخزّن في محرّك البحث ؟ هل يتمّ تخزين جميع المعلومات الموجودة في الموقع أم جزء منها وأيّ جزء ؟
- كيف تصنّف وتقيّم هذه المعلومات في قواعد البيانات ؟
- هل يدعم محرّك البحث إمكانيّات البحث المتقدّم (مثلاً استعمال صيغ بحث منطقيّة

معقّدة) ؟

بعد العثور على محرّك البحث الملائم لمجال البحث ، عليك أنْ تقرّر كيف تكتب الاستعلام ومتى يجب استعمال البحث البسيط أو البحث المتقدّم . هذه الاختيارات وغيرها من الكثير (ستوضّح في الفرع التالي) تهدف في هذه المرحلة إلى تقليص نتائج البحث إلى نتائج صحيحة ومطابقة أكثر لموضوع البحث [Vermel, 1999] .

تضييق مجال البحث :

في أغلب الأحيان ، عندما نكتب استعلامًا في محرّك البحث ، نحصل على مجموعة كبيرة جدًا من عناوين المواقع التي يجب أنْ يكون لها صلة بموضوع البحث ، حيث يرافق هذه العناوين في معظم محرّكات البحث وصف نصّي للموقع ، يمكن استغلاله لتحديد ما إذا كان الموقع يلبّي حاجتك أم لا . في الحقيقة لا نلاحظ دومًا ، بل في أغلب الحالات ، بأنّ هناك عناوين لمواقع ، لا تعرض معلومات عن موضوع البحث ، بل تحتوي على كلمة البحث في ترابط آخر ، يختلف كليًا عمّا نريد . لذلك ولتجنّب كثرة المواقع غير المتّصلة بموضوع البحث نحاول تقليص مجال البحث عن طريق استعمال البحث المتقدّم الذي يتيح استخدام صيغ بحث منطقيّة (بوليانيّة) معقّدة وأيضًا إشارات ورموز أخرى لنفس الغرض .

على سبيل المثال معظم محرّكات البحث تعطي إمكانيّة اختيار لغة المواقع ، حيث بإمكانك تحديد اللغة التي تريدها لكي تحصل على مواقع مكتوبة في هذه اللغة فقط وبذلك قمت بحصر نتائج البحث من البداية بشكل جيّد .

الجدول التالي يعرض عمليّات منطقيّة ، إشارات ورموز أخرى ، يمكن استعمالها من أجل كتابة الاستعلام بشكل أدقّ ، وبذلك الحصول على نتائج بحث ملائمة

ومطابقة أكثر . هذه القواعد معرفة في محرّك البحث Altavista (http:/ / www.altavista.com) ، لكنّها تستعمل في محرّكات بحث أخرى بشكل مشابه ومن أجل التأكّد من ذلك عليك تصفّح المساعدة الموجودة في محرّك البحث الذي تستعمله .

العمليّات المنطقيّة (البوليانيّة) :

كلمة دالّة	رمز	نتيجة	مثال
AND	&	ابحث عن المواقع التي تحتوي على جميع كلمات البحث	Internet AND Society
OR	|	ابحث عن المواقع التي تحتوي على كلمة واحدة على الأقلّ من كلمات البحث المذكورة	Woman OR female
NOT	!	اخرج جميع المواقع التي تحتوي على الكلمات المذكورة	Food AND NOT German
NEAR	~	ابحث عن المواقع التي تحتوي على كلمات قريبة من كلمة البحث	NEAR technology

إشارات أخرى :

علامة	نتيجة	مثال
+	تشترط العلامة + قبل كلمة البحث على ظهور الكلمة في الموقع .	+Arabic+Language
-	العلامة – قبل الكلمة تؤدّي إلى إظهار المواقع التي لا تحتوي على هذه الكلمة .	Action -horror
*	ابْحَثْ عن المواقع التي تحوي كلمة البحث وكلّ تكملة لها .	Child*: childhood, children
	تستعمل عند البحث عن جملة أو مصطلح مكوّن من عدْة كلمات . ويجب أن تكون جميعها في الموقع (وفي نفس الترتيب) .	"Artificial Intelligence"

ملاحظات :

1. شرط أن تكون لهذه الجامعات مواقع على شبكة الإنترنت .
2. هذه إمكانيّة واحدة ويمكن الوصول إلى نفس النتيجة بطرق أخرى .
3. يطلق عليه كثيراً اسم العنكبوت (Spider) أو الروبوت (Robot) .

ثبت المصادر والمراجع :

1. **[Sander-Beuermann, 1998]:** Sander-Beuermann, W., "Schatz-sucher", C't 13/1998, p. 178,
   http://www. heise. de/ct/98/13/178/.
2. **[Harris, 1997]:** Harris, R., "Internet search Tips and strategies", southern California college, version date: 31.031997;
   http://www. sccu. edu/faculty/R_harres/howlook.htm.
3. **[Vermel, 1999]:** Vermel, J. "Using search engines efficiently", Beitberl-College, 1999,
   http://ftp.beitberl.ac.il/~jean/fd/modules/modl/one_one/index.html.
4. **[4]:** "Die Suchfibel",
   http://www.suchfibel. de/2kunst/manuell.htm, 1999.
5. **[5]:** "Boolean Searching on the Internet", Univesity at Albany libraries,
   http://www.albany.edu/library/internet/boolean.html, 1999.
6. **[طيبي ،1999]:** طيبي مؤنس ، «شبكة الإنترنت للطالب والمعلم» . صدى التربية ، آذار 1999 ، المجلّد 47 ، صفحة 30 .
7. **[Haskin, 1998]:** D. Haskin ، «مواقع التفتيش الجبّارة» ، إنترنت وورلد ، العدد الأوّل ، كانون الثاني 1998 .