

الفصل السابع  
البحث في "الانترنت" المشكلات  
والأدوات

تحتوي انترنت، وشبكة (ويب) على وجه الخصوص حالياً على ملايين الصفحات من المعلومات معظمها على شكل نصوص. وتنمو كتلة النصوص التي تحتوي عليها انترنت بمعدلات متسارعة فضلاً عن خضوعها لعمليات التحديث المستمرة على مدار اليوم بل الساعة. ويطرح كل ذلك تحدياً جديداً أمام أولئك الذين يبحثون عما يحتاجونه من معلومات في هذا البحر الزاخر بالمعلومات.

غير أن مشكلة البحث عن المعلومات في النصوص لم تنشأ مع ظهور انترنت، بل تعود جذورها إلى عهد الكتب المطبوعة، ولتسهيل عملية البحث عن موضوع معين في كتاب مطبوع، تضم الكتب الحديثة في صفحاتها الأخيرة فهرساً للكلمات الدلالية الهامة (الكلمات المفتاحية) وأمام كل منها أرقام الصفحات التي وردت بها الكلمة. وقد إتمدت أساليب البحث عن المعلومات بواسطة الكمبيوتر على منهجية مشابهة لمنهجية استخدام كلمات دلالية (Keywords) كتلك المستخدمة في الكتب التقليدية. وتمتاز الأدلة المستخدمة في نظم إسترجاع المعلومات بإستخدام الحاسب، بإتساع نطاقها مقارنة بأدلة الكتب.

ومشكلات البحث عن المعلومات خلال الانترنت ناتجة من سببين رئيسيين، السبب الأول هو كثرة المعلومات الموجودة على الشبكة، فهناك ملايين صفحات الوب المتشابهة وكذلك مواقع الجوفر (Gopher) ومواقع (FTP) بالإضافة إلى الملايين من مجموعات الأخبار (News-Group) وقوائم البريد (Mailing Lists). السبب الثاني هو عدم وجود هيئة أو منظمة موحدة تشرف على تنظيم هذه المعلومات، بمعنى أنه لا يوجد كتالوج مركزي لهذه الموارد المتاحة ولا يوجد مكان واحد موحد يمكنك الذهاب إليه للبحث عما يريد.

ومع اتساع قاعدة النصوص. كما هو الحال مع شبكة انترنت، تصبح عملية إسترجاع المعلومات بإستخدام دليل الكلمات غير عملية وتؤدي هذه الطريقة إلى إسترجاع أعداد كبيرة من النصوص التي تحتوي على الكلمات الدلالية التي تبحث عنها، بالإضافة إلى أن هذه النصوص قد لا يكون لمعظمها صلة وثيقة بالموضوع قيد البحث على الرغم من إحتوائها على الكلمة المطلوبة

على سبيل المثال إذا كنت تبحث عن معلومات عن تاريخ بعض الدول وأدخلت أسم الصين. مثلاً، كمفتاح للبحث ستحصل على جميع الوثائق التي تحتوي على كلمة الصين(China) غير ان إحتواء بعض الوثائق على كلمة الصين لا يعني بالضرورة انها ستكون وثيقة تتعلق بتاريخ الصين، فعندما ترد الكلمة في وثائق من النوعيات أطباق صينية، عادات صينية، علاقة الصين باليابان، الحروف الصينية ... إلخ، وكما هو موضح فهي موضوعات ليست لها علاقة بالتاريخ الصيني القعديم.

ولمساعتك على الوصول إلى مصادر المعلومات المطلوبة وفرزها وتنقيتها من المعلومات غير المرتبطة بها وباستخدام أدوات للبحث، تستخدم تقنيات مثل "وزن الكلمة" (Word Weighing).

ووفقاً لهذه التقنية لا تعامل سائر كلمات النص على قدم المساواة، بل يسند إلى كل منها (وزن) يتناسب مع درجة أهمية الكلمة المعنية في النص فكلمة الصين الواردة في تقرير حول الصين كانت رقم (1) من قبل (China was No.1 before) مثلاً، سيكون لها وزن أكبر تاريخياً مما للكلمة ذاتها في سياق وثيقة تتضمن مجموعات الاختيار في الصين (Newsgroup in China). وتأتي نتائج البحث القائم على هذا المنهج مرتبة بحسب درجة أهميتها المتوقعة بالنسبة للباحث، ويؤدي الانتشار والنمو المتسارع للانترنت إلى إزدهار خدمات البحث عن المعلومات في داخل هذه الشبكة الكونية. وستعرض فيما لبعض المواقع التي توفر مثل هذه الخدمات مع عناوين هذه المواقع على الشبكة.

♦ أهم المواقع التي تقدم خدمات البحث في شركة انترنت:

1. اكسيت (Excite):

تقدم هذه الخدمة شركة (text SoftwareArchi) وقد تم تقديمها على الشبكة منذ وقت قصير . وتشمل هذه الخدمة أداتين ةتتيح البحث عن مواقع معينة على شبكة ويب (Wib) الأولى: البحث في الشبكة (Network) والبحث في المجموعات الاخبارية (wsgroupsNe) الخاصة بشبكة (Usenet). ويمكن استخدام كلمات مفتاحية في عملية البحث أو إستخدام مفاهيم كمنطلق للبحث مع

إمكانية تحديد نطاق للبحث. أما الأداة الثانية فهي مراجعة الشبكة (Netreviews) وهي توفر مخططاً هرمياً لمحتويات شبكة ويب ميوبة حسب الموضوعات وتفرعاتها.

ويتم تحديث قاعدة البيانات الخاصة باكسيت أسبوعياً وهي تضم حالي ١٠٥ بليون صفحة ويب و ٥٠٠٠٠٠ صفحة أخبار وعنوان هذه الخدمة على شبكة ويب هو: <http://www.excite.com>

### ٢. انفوسيك (Infoseek)

تتوفر هذه الخدمة بشكل مجاني أو مدفوع الأجر. وتغطي الخدمة المجانية شبكة ويب ويوزنت (Usnet). ونتيجة لذلك يتزاحم معظم المستخدمين على هذه الخدمة مما يصعب معه الوصول إليها. وتغطي هذه الخدمة المجانية ما يقرب من مائة عنوان للبحث أما الخدمة المدفوعة الأجر فتغطي ما يقرب من مائتين عنوان بالإضافة إلى سرعة استجابتها لطلبات البحث. وتقدم " انفوسيك " نتائج عملية البحث في صورة عناوين الموضوعات المرئية بحسب درجة الصلة المتوقعة لموضوعاتها مع مادة البحث.

وعنوان هذه الخدمة على شبكة ويب هو .

<http://www.Infoseek.com>

### ٣. ليكوسن (Lycos)

وتعتبر هذه الخدمة من أهم وأقوى أدوات البحث في شبكة انترنت. ولهذه الخدمة دليل للكلمات يضم أكثر من ثمانية ملايين كلمة تغطي أكثر من ٩٠% من محتويات شبكة ويب. ولهذه الخدمة مزايا فريدة تشمل إمكانية البحث عن النصوص والصور والمواد الصوتية والفيديو وقائمة تضم حوالي ٢٥٠ من أكثر مواقع شبكة ويب شعبية ويفضل استخدام هذه الشبكة في أعمال البحث المكثفة وليس البسيطة.

عنوان خدمة ليكوس على الشبكة هو :

[http:// www.Lycos.cmu.edu](http://www.Lycos.cmu.edu)



وزن " كل كلمة مقارنة مع الموضوع الرئيسي للنص وعنوان هذه الخدمة هو:

<http://www.opentext.com>

#### ٨. مفهرسات الويب ( Web Crawler ) :

تقوم شركة أمريكا أون لاين ( America Online ) بتوفير هذه الخدمة. لذلك فهي تعتبر خدمة مدفوعة الأجر وليست مجانية. ويستخدم فيها مايشبه أجهزة الروبوت في جمع الكلمات وفهرستها عبر ملايين الصفحات في شبكة ويب " ويب ". وتمتاز هذه الخدمة بسهولة الإستخدام وسرعة الإستجابة لطلبات البحث القائم على المحتوى. كما تحتوي الصفحات الخاصة بهذه الخدمة على معلومات إحصائية هابة عبر شبكة ويب وعنوان هذه الخدمة هو:

<http://www.woberawler.com>

#### ٩. البرمجيات ( Shareware ) :

تقوم هذه الخدمة أسرع وسيلة للحصول على البرمجيات من على الشبكة . ويمكن من خلالها البحث خلال ١٧٠,٠٠٠ ملف وعارضة ( Browsing ) بالإضافة إلى إمكانية تحميل هذه البرمجيات ( Download ) على حاسبك الشخصي باستخدام خدمة FTP للإستفادة منها وعنوان هذه الخدمة هو:

<http://www.Shareware.com>

#### ١٠. خدمات الوب ( ervers TouristWeb S ) :

يضم هذا الموقع دليلا شاملا للمواقع العلمية والأكاديمية المسجلة لدى الكونسورنيوم ويقوم الذي يتولى الإشراف على شبكة ويب . يضم هذا الدليل دليلا فرعيا من النوع الفهرسي ( rectoryWeb Servers Di ) مبوبا على أساس جغرافي ودليلا فرعيا آخر على هيئة خريطة للعالم ( ouristVirtual T ) يمكن التعامل معها بالضغط على مواقع معينة منها بالفأرة. وهذه الخدمة لا توفر آلية للبحث من النوع الذي توفره الخدمات الأخرى.

عنوان هذه الخدمة هو:

[http://www.w3.org/hypertext/Daa-Sources/  
www/Servers.html](http://www.w3.org/hypertext/Daa-Sources/www/Servers.html)

### ١١. المكتبة الإلكترونية ( *Electric Library* )

يمكن عن طريق إستعراض محتويات هذه المكتبة عدم الحاجة إلى البحث خلال الوب . فهي تضم قاعدة بيانات تحوى ألف صحيفة نصية ( ewspaperText N ) والعديد من المجلات والدوريات العلمية (rnalsAcademic Jou) بالإضافة إلى العديد من الصور والمراجع العلمية ( oksReference Bo ) والأبحاث والموضوعات الفنية...إلخ.

### ١٢. آلة البحث العملاقة ( *Search* )

يضم هذا الموقع ٢٥٠ آلة بحث يتم إستخدامها كآلة بحث عملاقة للغوص خلال أعماق الانترنت . وتتيح هذه الآلة الولوج إلى كل الخدمات الرئيسية التي تعرضها الشبكة ومئات من قواعد البيانات المتخصصة وعنوان هذه الخدمة هو:

<http://www.co.Search>

### ١٣. الكتالوج العام للانترنت ( *Whole Internet Catalog* )

يتيح هذا الكتالوج ١٦٠٠ موقع يمكن الولوج إليها مجاناً مع توصيف هذا الموقع بتقديم نبذة عن كل منها والدليل الذي توفره هذه الخدمة منظم بطريقة جيدة ولكنه لا يوفر آلية البحث من النوع الذي توفره الخدمات الأخرى وعنوان هذا الموقع هو:

<http://gnn.com /Wic/wics/index.html>

### ١٤. ياهو ( *Yahoo* )

هذا الموقع من أكثر المواقع شعبية على الانترنت من حيث المترددين عليه. وهذا أول موقع قدم خدمة البحث على الشبكة. لذلك فهو يحتوي على دليل منظم بطريقة جيدة ويضم كل محتويات شبكة ويب بحسب الموضوعات. ويمكن الولوج إلى هذه المحتويات بمتابعة إرشادات مرجعية تقود إلى الموضوعات ذات الصلة. ومن أهم المزايا التي تقدم، ها هذه الخدمة ميزة انشاء مواقع على الويب ( Web Lunch ) وادراجها في الدليل كما يتضمن دليلاً بمواقع ويب

الأكثر شعبية وقوائم بمواقع مميزة ومثيرة للإهتمام وينصح بزيارتها.  
وعنوان هذه الخدمة هو:

<http://www.Yahoo.com>

## ١٥. فهارس الانترنت (Net):

يوجد صفحات من الويب تسمى فهارس (Directories) وهي من أهم الصفحات التي تساعد في عملية البحث عن موضوع عام (General Topics) مثل الفنون و التعليم و الإدارة و الأنشطة و يحتوي الدليل على مواقع (Sites) مرتبة حسب الموضوع.

فمثلا يعرض دليل إكسيت (Excite) لمواقع تحوي موضوعات عديدة من بينها: الفنون - الأعمال - الحساب - التعليم - التسلية- الصحة والدواء - الأنشطة- المياه - المال و الاستثمار- الأخبار- الصفحات الشخصية - السياسة و القانون - الديانات- العلوم - التسوق - الرياضة.

و كما هو واضح تغطي هذه المواقع العديد من مجالات الاهتمام و تنتج عند الولوج إليها مصادر واسعة من المعلومات، و عند النقر عليها ستظهر لك قائمة بالموضوعات الفرعية المندرجة تحت الموضوع الأساسي، وهكذا تتفرع داخل القوائم حتى تصل إلى ما تبحث عنه، و فيما يلي نقدم بعض ادلة الانترنت الأكثر شعبية:

A2Z -Amazing Enviromental Organization Web  
Directory -Excite - Gamelan - Infoseek Guide - Magellan  
- Point - W3 Server - World Wide -Arts Resourees -  
Yahoo !

## و الجدول الآتي يعرض لعناوين هذه المواقع:

العنوان	الدليل
<a href="http://www.yahoo.com">http://www.yahoo.com</a>	Yahoo
<a href="http://theyellowpages.com/feedback.htm">http://theyellowpages.com/feedback.htm</a>	Yahoo pages.com
<a href="http://www.netcenter.com/netcenter/whats-new/whats-news.html">http://www.netcenter.com/netcenter/whats-new/whats-news.html</a>	NetCenter
<a href="http://www.Lib.umich.edu/chheme">http://www.Lib.umich.edu/chheme</a>	Internet

العنوان	الدليل
html	resourceGuid
<a href="http://galaxy.einet.net">http://galaxy.einet.net</a>	Galaxy
<a href="http://www.gold.net/gold/">http://www.gold.net/gold/</a>	The Global On-Line Mckinety Groups
<a href="http://www.Pointcom.com/">http://www.Pointcom.com/</a>	Magellan Point
<a href="http://www.Homecom.com/global\pointers.html">http://www.Homecom.com/global\pointers.html</a>	Communications Pointersto Pointers
<a href="http://www.stpt.com/until/submit.html">http://www.stpt.com/until/submit.html</a>	Starting Point
<a href="http://webSPACE.com.nvoyager">http://webSPACE.com.nvoyager</a>	Web Vayager
<a href="http://nearnnet.gnn.com\wie\index.htm">http://nearnnet.gnn.com\wie\index.htm</a>	The whole Internet cataiog

و من أفضل الطرق للوصول إلى هذه الأدلة هو الرجوع إلى أحد العناوين الآتية:

<http://home.netseape.com\home\internet-directory.html>

<http://www.lip.umich.edu\ehhome.html>

عند استخدام أحد هذه الأدلة مثل دليل اكسيت في عملية البحث، ومن خلال برنامج نيت سكيب ( Net Scap ) تظهر الموضوعات التي تحتوي على الكلمة التي تبحث عنها بجانبها نسبة تواجه الكلمة داخل الوثيقة و درجة الثقة ( Confidence ) و تعرض خدمة اكسيت إمكانية تعميق عملية البحث ( Refin Search ) أو إجراء بحث جديد ( New Search ) كذلك استعراض الأسئلة العامة المتعلقة بالموضوع بالإضافة الى العديد من الآليات التي تسهل عملية البحث.

و كل مواقع البحث و ما تقدمه من مصادر معلومات تضعك دائما في موقف المتساءل أياً من هذه المواقع يعطي أفضل المصادر ( Best Resour ) و أياص منها يعطي أكثر المصادر ( Most Resourees ).

و بالرغم من تأثير حجم قاعدة البيانات و جودة الفهرسة على نتائج البحث عنها، فان كيفية إدخالها ما زالت تؤثر على جودة النتائج، فليست كل آلات البحث تعامل الكلمة أو الجملة التي تدخل بنفس الطريقة.

على سبيل المثال يقوم ( Infoseek ) بالبحث عن الكلمة كلها أو أجزاء منها، فمثلا عند البحث عن كلمة ( Compression ) يمكنه مطابقة الجزء ( Compres ) و بمجرد أن يجده يضمه إلى نتائج البحث، بينما ليكوس ( Lycos ) بمطابقة الكلمة أو ما زاد عليها، فمثلا عند البحث عن كلمة ( Fluid ) يمكنه مطابقة ( Fluidity ) و ضمه إلى نتائج البحث.

و تسمح بعض خدمات البحث باستعمال الكارت العام ( Wild Cards ) مثل "؟"، "\*"، " " داخل الجملة مسار البحث. ومن الامثلة هذه الخدمات خدمة ( Aliweb ) و عنوانها:

<http://web.nexor.co.uk/public/aliweb.html>

و يمكنك في معظم آلات البحث استخدام العوامل المنطقية ( Logical Operators ) مثل AND أو OR أو NOT . و تتوفر انفوسيك ( Infoseek ) و فهرس النص المفتوح ( Open Text Index ) بإمكانية البحث عن الكلمات المتقاربة ( near ) أو المتوالية ( Followd Next ) من بين الكلمات المحددة في عبارة البحث.

#### ٩-٤ دمج أدوات البحث:

يمكنك استخدام آلية من آليات البحث المتعدد (MetaSearch)، و التي تتيح استخدام أكثر من آلة للبحث في وقت واحد، و من أشهر صفحات الويب التي تعرض هذه الخدمة المدمجة صفحة ( Savvy Search ) و عنوانها:

<http://www.es.colostate.edu/dreiling/smartform.html>

و صفحة Meta Crawler و عنوانها:

<http://www.es.washington.edu/research/projects/melaerawler/www/home.html>

و باستخدام أدوات البحث المدمجة لا يسمح لك باستخدام مميزات آلة البحث المنفردة مثل العوامل المنطقية و مثيلاتها مما قد يؤثر بالسلب على دقة البحث.

#### ٩-٥ سبعة نصائح مفيدة أثناء عملية البحث:

آلة البحث أو خدمة البحث تستخدم قاعدة بيانات عبارة عن فهرس يضم مجموعة من الكلمات أو العبارات يصاحبها عناوين لمواقع على شبكة انترنت. و عند إستخدامك لهذه الأدوات يجب أن تراعي جيدا أختيار و وضع الكلمات المناسبة لكي يتم مطابقتها و الحصول على فائدة أكبر من عملية البحث، لذلك إليك سبعة نصائح مفيدة لتوفير أقصى استفادة من أدوات البحث:

١. أختار كلمات بحث مميزة و ليست دارجة، و كلما كانت الكلمة التي تدخلها مميزة و غير معتادة و تعبر عن حاجتك كلما كانت النتائج أفضل و أدق و أسرع.

٢. لا تستخدم لغة التخاطب العادية: بالرغم من أن بعض أدوات البحث تتيح لك استخدام اللغة العادية و التي يمكنك من أن تسأل عما تبحث و كأنك تحدث أنسان مثلك لكن يفضل دائما استخدام الكلمات الهامة التي تعبر عن موضوع بحثك.

٣. إستخدام اكثر من آلة بحث: كرر تجربة البحث بإستخدام أكثر من آلة للتأكد من حصولك على افضل و أقرب النتائج.

٤. اقرأ تعليمات إستخدام آلة البحث: توفر في كل آلة مجموعة من التعليمات التي تساعدك في إجراء البحث بكفاءة مثل المعاملات (Operators) و المحددات (Delimiters) و قواعد البحث (Rules) لذلك يجب أن تقرأ جيدا هذه التعليمات قبل إستخدامك لآلة البحث.

٥. الكلمات ذات المقطعين: عند البحث عن كلمة ذات مقطعين مثل Turbo-Propeller حاول تجربة كل الأشكال الممكنة لكتابتها مثل:

Turbo-Propeller - Turbo-Propeller - Turbo-Propeller

٦. استخدام المعاملات المنطقية: تدعم بعض آليات البحث استخدام المعاملات المنطقية، و ينصح باستخدام المعامل (Not) إذا كان متاحا و الذي يمكنك من أسبعاد أحد من البحث: فمثلا عند البحث بالصيغة الآتية:

***Aircraft NOT Civil Subsonic***

فإنك تحصل على معلومات عن كل الطائرات و قد تم استبعاد المعلومات الخاصة بالطائرات المدنية.

٧. أحرص على ذكر البدائل و المرادفات: و قد يكون ما تبحث عنه موجود بإسم مرادف أو أسم بديل لذلك إحرص على إدخال معظم الكلمات التي تدل على ما تبحث عنه لكي يزيد فرصتك في الحصول عليه.