نظم المعلومات البيبليوجرافية الآلية

إعداد

د/زينب علي بكري علي

قسم المكتبات وتكنولوجيا المعلومات

كلية الآداب- قنا

العام الجامعي

2023/2024

محاضرات في

نظم المعلومات
البيليوجرافية الآلية

إعداد

الدكتورة/ زينب على بكرى على

كلية الآداب بقنا- جامعة جنوب الوادي

2024/2023م

بيانات الكتاب

الكلية	: الآداب
الفرقة	: الرابعة
التخصص	: المكتبات والمعلومات
تاريخ النشر	: 2024/2023
عدد الصفحات	: 158 صفحة
إعداد	: د. زينب علي بكري علي

بسم الله الرحمن الرحيم

" ((يَا أَيُّهَا الَّذِينَ آمَنُوا اتَّقُوا اللَّهَ وَقُولُوا قَوْلاً سَدِيداً * يُصْلِحْ لَكُمْ أَعْمَالَكُمْ وَيَغْفِرْ لَكُمْ ذُنُوبَكُمْ وَمَنْ يُطِعِ اللَّهَ وَرَسُولَهُ فَقَدْ فَازَ فَوْزاً عَظِيماً)) "

صدق الله العظيم .

[الأحزاب:70-71] .

قائمة المحتويات

مقدمة: 12

الفصل الأول: نظم استرجاع المعلومات 14

المفهوم؛النشأة؛ مراحل التصميم 14

مقدمة: 15

1/ مفهوم نظام المعلومات: 16

2/ أنواع نظم المعلومات: 17

3/ مفهوم استرجاع المعلومات: 17

4/ مفهوم نظم استرجاع المعلومات: 18

5/نظم استرجاع المعلومات: النشأة والتطور 20

6/ أشكال نظم استرجاع المعلومات: 22

7/ مكونات نظم استرجاع المعلومات: 22

8/ آلية بناء ومراحل تصميم نظم استرجاع المعلومات : 23

9/ الخصائص العامة لنظم استرجاع المعلومات الآلية: 35

10/ خدمات نظم استرجاع المعلومات الآلية: 36

11/ لغات استرجاع المعلومات: 37

12/ الجوانب المؤثرة في كفاءة نظم استرجاع المعلومات: 47

الفصل الثاني..... 51

تقييم نظم استرجاع المعلومات 51

مقدمة: 52
1/ مستويات تقييم النظم: 54
2/ العوامل والمقاييس النوعية: 56
3/ مقاييس الاستدعاء والدقة: 59
4 / مفاهيم الكفاءة: 65
5/ العوامل المؤثرة في الكفاءة: 70
6/ تأثير دور الوسيط على كفاءة الاسترجاع: 71
7/ دور المستفيد في كفاءة استرجاع المعلومات: 82
8/ اختيار قاعدة البيانات: 88

الفصل الثالث 93

قواعد البيانات: نموذج لنظم استرجاع المعلومات الآلية 93

مقدمة: 94
1/ قواعد البيانات: 95
1/1 البيانات(Data): *95*
2/1 المعلومات: *96*
3/1 تعريف قاعدة البيانات: *96*
4/1 مكونات قاعدة البيانات: *98*
5/1 أهمية وفوائد قواعد البيانات: *105*
6/1 أنواع قواعد البيانات (Types of Data Bases): *107*
7/1 نظم إدارة قواعد البيانات Database Management System: *112*
2/ مراحل تطور قواعد البيانات البيليوجرافية: 115
1/2 مرحلة الكشافات والبيليوجرافيات المطبوعة: *115*
2/2 مرحلة القواعد البيليوجرافية المحسبة على الخط المباشر عبر خط التليفون والمودم: *116*
3/2 مرحلة القواعد البيليوجرافية على الأقراص المدمجة: *118*
4/2 مرحلة القواعد البيليوجرافية المباشرة عبر الانترنت: *121*

الفصل الرابع 124

استراتيجيات البحث في قواعد البيانات 124

مقدمة: 125
1/ برمجية البحث والاسترجاع فى قاعدة البيانات: 126
2/ استراتيجية البحث: 127
3/مستويات استراتيجية البحث: 128
4/ آلية تنفيذ استراتيجية البحث: 130
5/ تقنيات البحث: 131

قائمة المصادر والمراجع 152

مقدمة:

تشهد احتياجات المستفيدين من المعلومات تعقدا في ظل تداخل التخصصات الموضوعية وتضخم الغنتاج الفكري الصادر في مختلف مجالات المعرفة، وقد أخذت المعارف في التداخل مع بعضها البعض مما أدى إلى تعقد التخصصات وظهور مجالات موضوعية جديدة نتيجة لتزاوج بعض حقول المعرفة مع بعضها الآخر، وأخذ الإنتاج الفكري يتضخم إلى الحد الذي شهدنا معه انفجار المعرفة، وأدت التقنية الحديثة بما تنطوي عليه من استخدام الحاسبات الآلية والاتصالات ونظم الشبكات–بما في ذلك شبكة الإنترنت–إلى تفاقم المشكلة، فالمعرفة في تزايد مستمر وهائل، والمستفيد لا يستطيع أن يحصر ما يحتاج إليه من معلومات منتشرة بلغات مختلفة وفي أماكن متعددة من ناشرين مختلفيين، ولا شك أن ذلك يجعل من الصعب على المستفيد الوصول إلى ما يحتاج إليه من معلومات دون أن تكون هناك أدوات مساعدة لتعريف المستفيد بأماكن نشر الإنتاج الفكري المتخصص في المجالات الموضوعية بمختلف لغاتها وأماكن نشرها

وتعد نظم استرجاع المعلومات بمثابة تلك الأدوات التي تنظم المعرفة وتساعد المستفيد في الوصول إليها واسترجاعها.

ويتناول مقررنا هذا نظم استرجاع المعلومات الآلية(نظم المعلومات الببليوجرافية الآلية)، يعرف بها ومراحل إنشائها ، ولغاتها ، وأساليب تقييمها، وكيفية البحث فيها.

ويتضمن الكتاب أربعة فصول على النحو التالي:

الفصل الأول: التعريف بنظم استرجاع المعلومات وتطورها والنظم الفرعية لها، ولغات الاسترجاع، وألية بناء نظم استرجاع المعلومات

الفصل الثاني: تقييم نظم استرجاع المعلومات، وبعض العوامل التي تعيق تحقيق النظم لأهدافها

الفصل الثالث: تعريف قواعد البيانات كنموذج لنظم استرجاع المعلومات، وأهميتها، ومراحل تطورها

الفصل الرابع: استراتيجيات البحث في قواعد البيانات

الفصل الأول: نظم استرجاع المعلومات

المفهوم؛النشأة؛ مراحل التصميم

مقدمة:

إن عملية إيجاد المصادر أو المعلومات المحتواه في المصادر لتلبية حاجة معينة هي عملية موجودة منذ القدم. ويبذل الإنسان كل جهده لجعل الوصول إلى المعلومات سهلا وميسرا. والإنتاج الفكري يتزايد بكميات رهيبة ويتعقد بصورة كبيرة، ولذلك أصبحت الطرق التقليدية للبحث أو استرجاع المعلومات غير فعالة وتستغرق وقتا طويلا. ومن ثم استنبطت أساليب حديثة للاسترجاع تسهل تداول وتوفير رصيد المعرفة المتزايد باستمرار. وتعتمد هذه الأساليب على أحدث ما توصل إليه العصر من تكنولوجيا.

فقد شكلت التطورات الكبيرة التي حدثت في مجال صناعة أجهزة الحواسيب سواء في أجزائها المادية hardware أو في مجال البرمجيات software دافعا قويا للعديد من مؤسسات المعلومات لاستثمار مميزات الحواسيب في بناء نظم استرجاع معلومات آلية. فبعد أن كانت نظم الاسترجاع التقليدية تعتمد على مصادر المعلومات الورقية المطبوعة في بناء مستودعاتها البحثية، وعلى الوسائل اليدوية في استرجاع المعلومات، أصبحت نظم الاسترجاع الآلية تعتمد على الوسائط القابلة للقراءة بواسطة الحاسوب. ونتناول هنا نظم المعلومات وقواعد البيانات باعتبارها من الأساليب الحديثة لاختزان المعلومات واسترجاعها ونقلها.

1/ مفهوم نظام المعلومات:

يرى بوركو أن نظام المعلومات هو إجراء منظم لتجميع وتجهيز واختزان واسترجاع المعلومات الموثقة لإرضاء حاجات متنوعة، وأن نظام المعلومات يتكون من مجموعة من المعلومات المسجلة، وأولى الأمر الذين يقومون بتنظيم المجموعة وحفظها وإجراء استرجاع، والمستفيدين الذين يطلبون المعلومات".

فهو إذا مجموعة من العناصر المتداخلة التي تعمل مع بعضها البعض بغرض جمع و معالجة و تخزين و توزيع المعلومات بغرض تلبية احتياجات المستفيدين و إرشادهم نحو ما يحتاجونه من أوعية معلومات".أو أنه عبارة عن مجموعة من النظم الفرعية التي تعمل مع بعضها البعض بهدف تحقيق أو إنجاز هدف معين.

وعموما فإنه ينبغي على نظام المعلومات أن يحقق ثلاثة متطلبات أساسية هي:

1) أن يكون قادرا على أن يعلم أو يخبر المستفيد أين يجد معلوماته.
2) ان يكون قادرا على نقل هذه المعلومات له عندما يقرر أنه يحتاجها.
3) أن يرد على أسئلة المستفيد في إطار حدود الوقت الذي يراه المستفيد مناسبا له.

2/ أنواع نظم المعلومات:

و توجد العديد من نظم المعلومات المعتمدة على الحاسب الالكتروني، وقد قسمها سالتون وماكجيل إلى خمسة أنواع هي:

1) نظم إدارة قواعد البيانات.
2) نظم المعلومات الإدارية.
3) نظم دعم اتخاذ القرار.
4) نظم الإجابة على الأسئلة
5) نظم استرجاع المعلومات .

وسوف يقتصر تناولنا هنا على نظم استرجاع المعلومات أو نظم استرجاع الوثائق.

3/ مفهوم استرجاع المعلومات:

يشير لانكستر في كتابه" أساسيات استرجاع المعلومات" إلى أن استرجاع المعلومات هو "النشاط الأساسي لما تقوم به مرافق المعلومات؛ من المكتبات ومراكز المعلومات والتوثيق، ومنتجي بنوك وقواعد البيانات، سواء المنشورة في شكل مطبوع أو الالكترونية، وأية نوعية أخرى من المرافق التي تقدم مصادر المعلومات لمجتمع المستفيدين".

وعلم استرجاع المعلومات هو "علم البحث عن الوثائق وعن المعلومات داخل الوثائق وعن الميتاداتا أو البيانات الوصفية المتعلقة بالوثائق بالإضافة إلى البحث في قواعد البيانات وشبكة الانترنت".

4/ مفهوم نظم استرجاع المعلومات:

وهي أحد النظم الفرعية في نظام المعلومات و وظيفته الأساسية هي اختزان واسترجاع المعلومات وفقاً لاحتياجات مجتمع المستفيدين، وسواء تم هذا داخل المكتبات أو خارجها فإنه مجرد وسيلة أو أداة يصل من خلالها الباحث والمستفيد إلى وعاء المعلومات الذى يبحث عنه، فالذي يُختزن عن هذه الأوعية بواسطة الحاسب الالكتروني في" بنك/نظام المعلومات الببليوجرافي" ليس إلا مجرد بيانات محدودة عن كل وعاء كعنوانه وتاريخه والمسئول عن محتواه الفكري ونشره وعدد صفحاته أو أوراقه ورؤوس الموضوعات فيه.

ويصف (فيكرى) نظم الاسترجاع الآلية بأنها مستودعات تعمل على اختزان مجموعة من الرسائل في إحدى الوسائل القابلة للقراءة بواسطة الحاسوب حيث يتم التعامل معها بواسطة مجموعة البرامج التي تعمل على إتاحة تلك الرسائل واسترجاع المناسب منها.

وتعرف الموسوعة العربية لمصطلحات المكتبات والمعلومات والحاسبات نظام استرجاع المعلومات بأنه "مجموعة من الإجراءات المميكنة عادة تستخدم في الرجوع إلى البيانات التي تحويها الوثائق وتكشيف تلك البيانات واختزانها بطريقة يمكن استعادتها عند الطلب".

كما يعرفه الصوينع بأنه "النظام الذى يتعامل مع معلومات وسيطة بين المستفيد والمعلومات النهائية التي يسعى إليها، فهو نظام لا يزيد

من الحالة المعرفية للمستفيد منه، ولكنه يحيط المستفيد علما بوجود معلومات عن الوثائق التى تتعلق بموضوع استفساره".

ونخلص من ذلك أن نظم استرجاع المعلومات يمكن تعريفها بأنها هي "النظم التي تسمح بتخزين وإدارة ومعالجة الوثائق بطريقة تسمح للمستخدم باسترجاع الوثائق التي تتوافق مع حاجته من المعلومات. وبمعنى أخر إنها مستودعات آلية لخزن المعلومات من مصادرها الأصلية بعد تحويلها إلى بيانات قابلة للخزن في وسائط مناسبة في الغالب تكون أقراصا ممغنطة أو صلبة أو ليزرية مع إمكانية قراءتها آليا بواسطة أجهزة الحاسوب ليتم الوصول إلى محتوياتها واسترجاع المعلومات المناسبة باستخدام مجموعة من البرامج القادرة على تهيئة استراتيجيات بحث تحقق نوعا من العلاقة الموضوعية بين المصطلحات المختلفة بهدف الوصول إلى استرجاع معلومات بعينها والتقليل إلى أدنى حد من استرجاع المعلومات غير المتوافقة".

ويمكن فهم نظم استرجاع المعلومات جيدا إذا تذكرنا أن المعلومات المعالجة تتكون من الوثائق. وفي ذلك السياق فإن استرجاع المعلومات يتعلق بالتمثيل والاختزان والوصول للوثائق أو ممثلاتها(بدائل الوثائق). والمعلومات المدخلة قد تشمل النصوص المكتملة للوثائق أو مقتبسات أو مستخلصات للوثائق. ومخرجات نظام استرجاع المعلومات إجابة لطلب بحث ما تتكون من مجموعة من الإشارات الببليوجرافية، والقصد

من هذه الإشارات الببليوجرافية هو إمداد المستفيد من النظام بمعلومات عن المواد أو الوثائق ذات الأهمية بالنسبة له.

5/نظم استرجاع المعلومات: النشأة والتطور

بدأت تظهر نظم استرجاع المعلومات في نهاية خمسينات القرن العشرين الميلادي ، وكان الهدف الأساس من نشأتها هو محاولة استثمار إمكانيات الحاسب الآلي في تحسين عمليات الطباعة والنشر ، والحاجة إلى معالجة التسجيلات الببليوجرافية في شكل قابل للقراءة بواسطة الآلات لأغراض مراجعة الأخطاء، والفرز، واستخدام الحاسب في عمليات الطباعة ، هذا فضلاً عن استخدام الأشرطة القابلة للقراءة عن طريق الآلات الناتجة عن هذا النشاط في إنتاج المزيد من النسخ وتقديم المزيد من الخدمات وبخاصة خدمة البحث عند الطلب . ومع أنه ليس هناك معلومات محددة ودقيقة عن معرفة تاريخ أول نظام استرجاع آلي، إلا أن هناك تواريخ محددة تعد نقاط هامة في تاريخ نشأة هذه النظم وتطورها، ومن أهم هذه التواريخ .

1959م : ظهور أول نظام استرجاع آلي في مختبرات مدفعية الأسطول في سلفرسبرنج في ولاية ماريلاند الأمريكية .

1960: ظهور نظام للاسترجاع الآلي أعدته جامعة وسترن رسيرف لحساب الجمعية الأمريكية للمعادن .

1962 : ظهور نظم استرجاع آلية أنشأتها وكالة المعلومات التقنية للقوات المسلحة الأمريكية والإدارة القومية للملاحة الجوية والفضاء (ناسا)

1963 : ظهور نظام استرجاع وتحليل الإنتاج الفكري في الطب والأحياء، المعروف باسم " مدلرز " ، الذي أعدته المكتبة القومية للطب في الولايات المتحدة الأمريكية .(تطوير للكشاف الطبي المطبوع)

1965: ظهور نظام الاسترجاع الببليوجرافي المعروف بنظام " دايالوج " لشركة لوكهيد.

ثم توالت نظم الاسترجاع في الظهور بعد ذلك، وعبر مسيرتها مرت بمراحل تميزت كل منها بسمات وخصائص محددة وفقاً للتطورات التقنية التي كانت سائدة في ذلك الوقت، والتي يمكن حصرها في أربع مراحل متميزة على النحو التالي :

المرحلة الأولى : نظم التجهيز على دفعات

المرحلة الثانية : نظم التجهيز والاسترجاع على الخط المباشر

المرحلة الثالثة : قواعد البيانات المحملة على الاسطوانات الضوئية المدمجة

المرحلة الرابعة : قواعد البيانات المتاحة عبر الإنترنت

6/ أشكال نظم استرجاع المعلومات:

تصمم النظم في العادة لاسترجاع المعلومات ذاتها أو للإشارة إليها في أي وعاء وثائقي تتواجد فيه تلبية لإجابة محددة عن تساؤل معين،ينبغي أن نفرق بين عملية استرجاع المعلومات وعملية استرجاع الوثائق؛ ويقصد بالأولى البحث في الإنتاج الفكري للحصول على وثيقة أو مجموعة وثائق حول موضوع معين، أما الثانية فتعني البحث عن وثيقة بعينها يعرف المستفيد عنوانها أو مؤلفها، وقد قسم لانكستر نظم استرجاع المعلومات إلى ثلاث فئات على النحو التالي:

- **نظم استرجاع البيانات:**

تعنى باسترجاع التسجيلات الببليوجرافية للوثائق.

- **نظم استرجاع الحقائق fact retrieval system**

وهى تلك النظم التي يرجع إليها الباحث للحصول على المعلومات بشكل مباشر أي النظم التي تعنى بالرد على الاستفسارات، كما في الخدمة المرجعية والخدمات التي تقدم في الشكل المقروء آليا مثل بنوك المعلومات.

- **نظم استرجاع الوثائق document retrieval systems:**

وهى النظم التي تختزن النصوص الكاملة لمجموعة وثائق يتم استرجاعها من خلالها.

7/ مكونات نظم استرجاع المعلومات:

يتكون نظام استرجاع المعلومات من ستة نظم فرعية أساسية هي:

- ❖ النظام الفرعي لاختيار الوثائق.
- ❖ النظام الفرعي للتكشيف.
- ❖ النظام الفرعي للغة.
- ❖ النظام الفرعي للبحث.
- ❖ النظام الفرعي للتفاعل بين المستفيد والنظام .
- ❖ النظام الفرعي للمضاهاة، والذي يقوم بمضاهاة تسجيلات الوثائق مقابل تسجيلات الاستفسارات.

وتتفاعل هذه النظم مع بعضها البعض لتحقيق الهدف من إنشاء النظام وهو استرجاع المعلومات. ويلاحظ أن النظم الخمسة الأولى، تعد من أنشطة فكرية يقوم بها البشر، في حين يعد النظام السادس نشاطا آليا، وعلى ذلك يرى البعض أن الإسهام الحقيقي للحاسب الآلي يكمن في عملية المضاهاة، حيث يعد وسيلة مضاهاة عملاقة.

8/ آلية بناء ومراحل تصميم نظم استرجاع المعلومات :

يمر نظام استرجاع المعلومات عند تصميمه بمجموعة من المراحل تبدأ بتحليل طلبات النظام، تصميم النظام، وتطبيقه وتقييمه ، فيما يلي عرض لهذه المراحل بمزيد من التفاصيل:

1/ المرحلة الأولى : تحليل متطلبات النظام.

في هذه المرحلة يتم إجراء دراسة جدوى لمعرفة مدى النجاح الذي يمكن تحقيقه من تطبيق النظام، بالإضافة إلى تحديد المقومات اللازمة للنظام مثل؛ المقومات البشرية والمادية والتكنولوجية والمالية والتنظيمية، وتحديد متطلبات المستفيد.

1/1 دراسة الجدوى:

تبدأ المرحلة الأولى بإجراء دراسة جدوى التي تعني معرفة الجدوى التقنية والاقتصادية والتنظيمية لإنشاء النظام الجديد، ويجب أن يقدم نشاط دراسة الجدوى إجابات على الأسئلة التالية:

1. ما المشكلات التي يفترض حلها عن طريق النظام المقترح؟
2. كيف سيحل النظام المقترح هذه المشكلات ؟
3. هل تقنيات المطلوبة للنظام متوافرة ومتاحة (الجدوى التقنية) ؟
4. كيف سيتم قبول النظام بواسطة العاملين المتعددين ذوي الميول والاتجاهات المختلفة–مثل ظاهرة مقاومة التغيير –(الجدوي التنظيمية)؟
5. هل النظام المقترح ذو مبررات اقتصادية–أ ي هل العائد يفوق التكلفة–(الجدوى الاقتصادية) ؟

وتمر دراسة الجدوى بعدد من الخطوات التالية:

- تحديد أهداف ومجال دراسة الجدوى : وتتمثل في توصيف الأهداف وصياغتها بأسلوب كمي، وتحديد القيود الداخلية والخارجية، معايير

الأداء– مثل التكلفة، الزمن، الدقة، الجودة ...– سياسة الإدارة– مركزية أو لا مركزية–، مدي مشاركة الوحدات التنظيمية.

- دراسة الموقف الحالي : مثل سياسات المكتبة ، ومعرفة تدفقات البيانات والمعلومات، وتحديد الإجراءات والعمليات الحالية وتعريف المشكلات التي يجب حلها.
- تحليل متطلبات المعلومات : تحليل البيانات ، تحليل القرارات.
- تحليل الحلول البديلة : الجدوى التقنية ، الاقتصادية ، التنظيمية.

2/1 تحديد المقومات اللازمة للنظام:

وتتمثل في المقومات البشرية، المادية، التكنولوجية، المالية، التنظيمية ، وهي علي النحو التالي:

1/2/1 المقومات البشرية:

والمقصود هنا الأفراد الذين سيعملون على النظام المقترح، ويلزم هنا التعرف إلى المهارات التقنية المتوافرة لديهم؛ لتحديد مدي قدراتهم على العمل على النظام الجديد وما إذا كان هناك حاجة للتدريب.

2/2/1 المقومات المادية:

وتتمثل في مصادر المعلومات الرقمية، حيث يتم تحديد هل سيتم توفير الوثائق كاملة أم يتم تحديد البيانات الببليوجرافية فقط ، وفي حالة الاختيار الأول؛ يتم تحديد أنواع المصادر المستهدفة لتحويلها إلى الشكل الرقمي وتحديد أشكال هذه المصادر من حيث الملفات التي ستتاح فيها مثل. DOC,XLS,PDF,HTML

3/2/1 المقومات التقنية (التكنولوجية):

وتشمل الأجهزة وبرنامج إدارة قواعد البيانات وشبكة المعلومات الداخلية LAN، وخطوط الاتصال بالإنترنت وهل سيتم الاتصال بتقنية DSLأم استخدام التقنيات اللاسلكية، كذلك تحديد السعة اللازمة لتخزين مصادر المعلومات الرقمية وسرعة المعالجات اللازمة للتعامل مع البيانات المختلفة في أثناء التشغيل (مواصفات الأجهزة) Hardware specifications)

4/2/1 المقومات المالية:

وهي التكاليف اللازمة لتطبيق النظام ومدي توافر الموارد المالية لذلك.

5/2/1 المقومات التنظيمية:

وتتمثل في السياسات والإجراءات الخاصة بتطبيق النظام المقترح ، من تحديد الوثائق التي سيتم تغطيتها والمجال الموضوعي للنظام، عملية الفهرسة والتصنيف والتكشيف، وإجراءات التوثيق الخاصة بتفسير البيانات المرمزة التي سيتم إدخالها إلى النظام.

3/1 تحليل متطلبات المستفيد:

تجري عملية تحليل النظم بغرض ايجاد طريقة لتعديل النظام الحالي أو تطبيق نظام جديد بشكل يلائم متطلبات المستفيد، وفي هذه الخطوة يتم التعرف إلى احتياجات ومتطلبات المستفيدين ؛ لان ذلك يساعد على تصميم نظام يعمل على تلبية الاحتياجات المعلوماتية لمستفيدين بشكل مناسب.

2/المرحلة الثانية : تصميم النظام:

في هذه المرحلة يتم تصميم عناصر المدخلات وعمليات المعالجة وتحديد طرق عرض المخرجات وإنشاء واجهة المستخدم user interface.

1/2 تحديد المدخلات: inputs

ينبغي أن يتم تصميم المدخلات وفقا لمعيار marc21حتى يتوافق النظام المنشأ مع الأنظمة الأخرى، مما يساعد علي التبادل الإلكتروني للبيانات، وتشتمل عملية تصميم المدخلات علي الخطوات التالية:

- يتم تحديد نظام إدارة قواعد البيانات المراد استخدامه، حيث يتوافر كثير من هذه البرمجيات مثل sql server, oracle، وهناك أيضا برنامج winisis الذي أصدرته منظمة اليونسكو عام 2003 م، ويتميز البرنامج الأخير بسهولة استخدامه في إنشاء وإدارة قواعد البيانات وإتاحتها على الخط المباشر online.
- إنشاء ملف البيانات الرئيسي main file الذي يتم فيه إدخال وتخزين البيانات الببليوجرافية مصادر معلومات.
- يتم تحديد الحقول المستخدمة المتمثلة في : المؤلف، العنوان، الموضوع، الكلمات المفتاحية ... وكذلك تحديد روابط Links في حالة الرغبة في استرجاع النص الكامل لمصادر المعلومات.
- يتم إعطاء رقم يكون الرقم هنا وفقا لمعيار مارك 21 مثل 100 مدخل رئيس لمؤلف فرد وهكذا، واسم لكل حقل وتحديد نوع

البيانات data type التي سوف تخزن به (حروف، أرقام، مختلط).وكذلك السماح بتكرار الحقول أم لا وإنشاء الحقول الفرعية وتحديد الحقول التي ينبغي أن تشتمل على بيانات والحقول التي من الممكن تركها خالية مع إنشاء رسائل messages تظهر لمستخدم النظام ؛ تعلمه بضرورة عدم ترك هذا الحقل خاليا أو أن البيانات المدخلة لا تتطابق مع نوع البيانات الخاصة بهذه الحقل.

- إعداد قائمة الاستبعاد stop word list؛ نظرا لأنه يوجد كثير من الكلمات الشائعة الاستخدام في الكثير من الوثائق مثل (حروف الجر والعطف وأل التعريف، وأسماء الإشارة) يتم إعداد قائمة بهذه الكلمات تسمى قائمة الاستبعاد، تشتمل على الكلمات التي يجب على النظام أن يمنع وصولها إلى الملف المقلوب (الكشاف) حتى لا تؤثر في دقة الاستدعاء. وجدير بالذكر أن الكلمات الموقوفة أو المستبعدة تختلف من نظام لآخر تبعا لطبيعة التخصص الموضوعي للنظام على سبيل المثال؛ نظام متخصص في المعلومات الطبية تختلف قائمة الاستبعاد فيه عن نظام في المعلومات المالية .
- إنشاء الملف المقلوب inverted file أو الكشاف عن طريق تحديد الحقول المراد الاعتماد عليها في استرجاع المعلومات من قبل المستفيدين مثل؛ حقول (المؤلف، العنوان، الموضوع)؛ حيث تساعد تلك الكشافات بسهولة في استرجاع المعلومات وهذه

الكشافات عبارة عن جداول بالحقول وتسجل القيم الموجودة بهذه الحقول وترتبها هجائيا مع تحديد عدد المرات فتكرار المصطلح في النظام وأرقام التسجيلات الموجود بها هذا المصطلح –على سبيل المثال– عند الرغبة في إنشاء كشافات لحقول المؤلف والعنوان والكلمات المفتاحية يتم إنشاء ثلاث كشافات.

- بعد أن يتم تحديد الحقول المراد استخدمها في عمليات الاسترجاع يتم تحديد نوعية التكشيف بمعنى؛ هل يتم التكشف للنص أي لجميع الكلمات الواردة بالحقل أم كلمة–كلمة، وأن كان التكشيف كلمة–كلمة أفضل؛ لأنه يساعد المستفيد علي استرجاع المعلومات في حالة إذا ما كان مثلا لا يعلم عنوان أو اسم مؤلف المصدر كاملا.

2/2 عملية المعالجةprocessing :

في هذه الخطوة يتم كتابة الكود الخاص باسترجاع المعلومات من النظام للعناصرالسابقة التحديد، وتعد لغة الاستعلام المهيكلة structured query language sqlهي الأداة الأساسية في هذه الخطوة حيث يتم استخدام العمليات التالية:

1/2/2 أوامر إنشاء استدعاء البيانات:

select : يستخدم هذا الأمر عند الرغبة في استرجاع البيانات.

update : لتحديث البيانات بالتسجيلات.

insert : لإنشاء تسجيلات جديدة.

delete : لحذف تسجيلات من النظام.

2/2/2 المتغير variables :

وهي المسئولة عن تخزين قيم البيانات التي يدخلها المستفيد للنظام، حيث يتم أخذ هذه القيم values ومضاهاتها بالقيم الموجودة بالتسجيلات الببليوجرافية في النظام، وتجدر الإشارة إلى أن هذه المتغيرات يتم حذفها من النظام بمجرد الانتهاء من الاستعلام المقدم من المستفيد ، حيث يتم إنشاؤها مرة أخرى عند التقدم بطلب استفسار جديد .

3/2/2 الشروط condition :

مثل (if..then,else) وذلك في حالة احتمال وجود شرط معين يتم تنفيذ مجموعة من الأوامر، وكذلك هناك selectcase وتستخدم بالتبادل مع القاعدة السابقة.

3/2 المخرجات outputs :

ويتم هنا تحديد أن الشكل المناسب لواجهة المستخدم والعناصر المكونة لها بحيث تسمح للمستفيد بسهولة الاستخدام والبحث واسترجاع المعلومات وذلك عن طريق القيام بالآتي:

- تحديد العناصر التي سوف يستخدمها المستفيد في البحث والاسترجاع مثل (المؤلف–العنوان–الموضوع،...)
- تحديد العوامل المنطقية boolean operators والتي تستخدم للربط بين الكلمات مثل(or–and–not)

- تحديد عدد التسجيلات التي سوف يتم عرضها في الصفحة الواحدة (10 تسجيلات مثلا)
- بناء أدوات مساعدة ترشد المستفيد وتوجهه إلى كيفية الاستخدام الأمثل للنظام.
- إتاحة كشافات بالمصطلحات الموجودة بالنظام مرتبة هجائياً حتى يتمكن المستفيد من التعرف إلى المصادر الموجودة تحت هذه المصطلحات وتعرف هذه الطريقة بالاسترجاع عن طريق التصفح Information Retrieval By Browsing.
- تحديد الطريقة التي يتم بها عرض النتائج ؛ مثل هل يتم عرض النتائج بشكل مختصر لكل تسجيلة وعند اختيار المستفيد لها يتم إظهار البيانات الكاملة.
- تحديد الأسلوب الذي يعرض به نتائج البحث مثل أن تعرض النتائج مرتبة هجائيا بالعنوان أم بالمؤلف أم بتاريخ النشر مثلا.
- وجود إشارات تساعد المستفيد على استخدام النظام وتوضح له كيفية إجراء عمليات البحث.

وتجدر الإشارة إلى أن الخطوات السابقة – المدخلات، المعالجة، المخرجات – تتم بشكل متواز، حيث إن تحديد شكل المخرجات يتطلب تحديد عناصر المدخلات وكذلك عمليات المعالجة التي سوف تتم على هذه العناصر، وتعد هذه المرحلة من أهم المراحل التي تمر بها نظم

استرجاع المعلومات وبالتالي ينبغي أن تتسم بالدقة والتكامل والحرص في عمليات التصميم الثلاث.

كذلك فإن مرحلة التصميم يتم فيها تحديد الصلاحية الخاصة بكل مستخدم أي العاملين بالمكتبة ، والعمل على تأمين النظام باستخدام كلمات المرور Passwords يتم تغييرها بصفة دورية، وكذلك تسجيل العمليات المختلفة التي تتم على النظام من قبل العاملين مثل إضافة أو حذف أو تعديل تسجيلات.

3/ المرحلة الثالثة : تشغيل واختبار النظام.

يتطلب تنفيذ النظام وتشغيله وجود خطة تنفيذية يشارك فيها محللو النظام والسلطات التنفيذية من مديري الإدارات المستفيدة من النظام، ويتم هنا تحديد طرق التنفيذ واختيار وتدريب طاقم العمل، كذلك يتم اختبار المدخلات Inputs Test لتحديد ما إذا كانت المستندات المتنوعة للمدخلات تطابق قواعد التصميم أم لا ؟ وهل يتم استيفاؤها بطريقة صحيحة من قبل مستخدمي النظام ؟ وكذلك اختبار المعالجة وإجراءاتها التي تتم داخل النظام، حيث إنه من الممكن أن يبدو النظام أنه يعمل بطريقة صحيحة؛ لأنه يعالج المدخلات وينتج ما يبدو مخرجا صحيحا، أو ربما ينتج مخرجات صحيحة لكنها تؤدي مهام إضافية غير مطلوبة.

وفي هذه الخطوة يقوم مصممو النظام بمراجعة الكود الخاص بعمليات المفردات والتأكد من سلامة وصحة عملية إدخال البيانات وشاشات المخرجات.

4/ المرحلة الرابعة : تقييم النظام.

تتمثل المعايير المستخدمة في تقييم نظم استرجاع المعلومات في تقييم التكلفة المادية والمعنوية للمستفيد ، وزمن الاستجابة، ومدى جودة المخرجات، وعائد التكلفة من استخدام النظام.

ويفصل الباحث العناصر السابقة على النحو التالي:

- **معايير قياس التكلفة للمستفيد** : يتم هنا قياس التكلفة المادية والمعنوية للمستفيد كما يلي:

- **التكلفة المادية** : وهنا يتم قياس التكلفة المادية (المالية) بالنسبة للمستفيد نظير كل عملية بحث يقوم بها وتشمل عناصر التقييم الآتي:

1) تكلفة الوحدة لكل إشارة وراقية (بيانات ببليوجرافية) متصلة بالموضوع يتم استرجاعها.
2) تكلفة الوحدة لكل إشارة وراقية جديدة لم تكن معروفة للمستفيد وتتصل بالموضوع وتم استرجاعها
3) تكلفة الوحدة لكل وثيقة متصلة بالموضوع تم استرجاعها.

- **التكلفة المعنوية** : والمقصود هنا قياس الوقت المرتبط بجهد المستفيد في إجراء البحث وذلك في حالة البحث غير المفوض. أما

في البحث المفوض فإنه يقاس بمقدار الوقت المرتبط بجهد المستفيد في التفاوض حول الطلب مع النظام ومقدار الوقت اللازم بعد الحصول على النتائج للتحقق من مدى صلة الوثائق المسترجعة بموضوع البحث ، وهنا يتم قياس العناصر التالية:

1) الجهد اللازم من قبل المستفيد في تعلم كيفية الإفادة من النظام.
2) الجهد اللازم في الإفادة الفعلية.
3) الجهد اللازم لاسترجاع الوثائق من النظام.

- **معايير قياس زمن الاستجابة** : ويمثل زمن الاستجابة في البحث المفوض الوقت الفاصل بين تقديم الطلب من جانب المستفيد وتلقي نتائج البحث. أما في البحث غير المفوض فهو الوقت المستغرق في إجراء البحث ، وتشمل العناصر التي يمكن قياسها هنا الآتي:

1) مدى الوقت المستغرق ما بين تقديم الطلب واسترجاع الإشارات الوراقية.
2) مدى الوقت المستغرق ما بين تقديم الطلب واسترجاع الوثائق.
3) وقت الانتظار لاستخدام نظام الخط المباشر.

- **معايير جودة المخرجات** : وتتمثل في قياس مدى تغطية النظام للموضوعات المختلفة ومدى حداثة المخرجات واكتمال البيانات الببليوجرافية ودقتها ، كذلك يتم قياس نسبة الاستدعاء والدقة، وتعبر نسب كل من الاستدعاء والتحقيق على مدى قدرة النظام على

التنقية Filtering أي قدرته علي استرجاع المطلوب وحجب ما سواه , وعموما فإنه يمكن عرض العناصر الخاصة بتقييم جودة المخرجات كما يلي:

1) مدى التغطية الموضوعية لنظام استرجاع المعلومات.
2) مداها اكتمال المخرجات (الاستدعاء)
3) ارتباط المخرجات بموضوع البحث (التحقيق)
4) حداثة المخرجات.
5) اكتمال البيانات الببليوجرافية للوثائق.
6) شكل المخرجات التي يقدمها النظام.

وأخيرا فهناك تقييم متصل بقياس عائد التكلفة والذي يحدد مدى أهمية النظام في مقابل تكاليف تشغيله والإفادة منه.

ويجب أن يتم التقييم بصفة دورية منتظمة وباستخدام المعايير والأدوات والمقاييس المناسبة لموضوع التقييم والاستفادة من نتائج هذا التقييم في تطوير النظام الحالي والعمل على التخلص من العيوب الموجودة به أو على الأقل تخفيضها إلى الحد الأدنى ، كذلك فإن التقييم ينبغي أن يضم كلا من النظام ومستخدميه.

9/ الخصائص العامة لنظم استرجاع المعلومات الآلية:

تتميز نظم استرجاع المعلومات بالعديد من الخصائص الواجب توافرها لفاعلية وكفاءة النظام، فمن أهم هذه الصفات:

1. المرونة بحيث يمكن الإضافة أو الحذف أو التعديل.
2. ترابط مكونات وعناصر النظام.
3. أن يكون هنالك توازن بين تكلفة النظام ومدى الاستفادة منه.
4. استخدام التقنيات الحديثة فى التنظيم والتحليل وبث المعلومات.

10/ خدمات نظم استرجاع المعلومات الآلية:

من أهم الخدمات التى تقدمها نظم استرجاع المعلومات المستخدمة فى مؤسسات المعلومات (مكتبات، مراكز معلومات،...) وهى:

1. استعراض مقتنيات قواعد البيانات المختلفة بحثا عن المعلومات المطلوبة ومصادرها بدلا من مقتنيات مكتبة واحدة أو مجموعة من المكتبات المحلية المتعاونة.
2. رفع كفاءة أداء العمليات الفنية فى مؤسسات المعلومات مع إنجازها فى أقل وقت وبأقل تكلفة.
3. إمكانية تبادل المعلومات مع مختلف المكتبات والحصول عليها بكل يسر.
4. إمكانية التنقل إلكترونيا من قاعدة بيانات لأخرى بحثا على المعلومات بغض النظر عن مكان وجودها.
5. حل مشكلة متابعة التزويد والاتصال بالناشرين بشكل آلي طلبا للكتب والدوريات وغيرها من مصادر المعلومات.

6. تسهيل مهمة أدوات البحث العلمي من ببليوجرافيا وكشافات ومستخلصات وفهارس.

11/ لغات استرجاع المعلومات:

إن للغة المستخدمة في البحث دور كبير في عملية الاسترجاع فمثلا لو كان هناك شخصان وكلاهما يبحثان عن الموضوع نفسه ولكن النتائج المسترجعة لكليهما مختلفة نوعا ما فذلك يعود إلى أن كلا منهما قد استخدم كلمات مختلفة للبحث.

وفي مجال استرجاع المعلومات يعرب المستفيدون عن حاجاتهم إلى المعلومات للوسطاء بمصطلحات لغوية، ويترجم الوسطاء هذه المصطلحات إلى لغة نظام استرجاع المعلومات، ولتوفير مقومات الاسترجاع الموضوعي تقوم النظم الحديثة لاسترجاع المعلومات بتجهيز الوثائق والطلبات بطرق تكفل مضاهاة الكلمات والعبارات الواردة في الوثائق بتلك الواردة في الطلب والمقصود بالمضاهاة هنا إجراء مطابقة بين المفردات التي استخدمها المستفيد للتعبير عن حاجته إلى المعلومات مع تلك التي تعكس المحتوى الموضوعي للوثائق المخزن في قواعد البيانات، لكن من المهم التأكيد هنا إلى أن المصطلحات المستعملة في اللغة الطبيعية في نظم خزن واسترجاع المعلومات في الغالب ليست متماثلة مع تلك المصطلحات التي استعملها المؤلف، ولا تلك التي يحتمل أن يستعملها المستفيد وتتداخل عوامل عدة في هذا التباين والتي يمكن إيجازها بالآتي:

أ- أن التطورات العلمية والتقنية غالبا ما تؤدي إلى ظهور مصطلحات جديدة باستمرار أو تؤدي إلى تغير مصطلحات ومفاهيم قديمة.

ب- أن مستوى الدقة والتحديد في بناء المفاهيم قد يختلف في لغة المؤلف عنه في لغة المستفيد.

ت- قد يلتقط المكشف المصطلحات من الوثيقة نفسها للتعبير عن محتواها الموضوعي أو يجتهد في اختيار تلك المصطلحات من خارج المحتوى النصي للوثيقة .

وتعتمد آلية عمل نظم استرجاع المعلومات على استخدام مجموعة من المصطلحات الموضوعية ليتم بواسطتها استجواب هذه النظم، والهدف الرئيسي هنا هو في تحقيق استجابة باسترجاع مجموعة من التسجيلات (الوثائق) التي يتطابق محتواها الموضوعي مع تلك المصطلحات. ولقد كانت رغبة المستفيدين في تحقيق التطابق التام والتقليل إلى أدنى حد ممكن من استرجاع تسجيلات غير متوافقة مع مطالبهم البحثية إلى التفكير في استخدام أساليب وطرق متنوعة، الغاية منها ضبط المصطلحات والمفردات بالطريقة التي تحد من التباين بين المصطلحات التي استخدمها المؤلف، مع المصطلحات التي تم تخزينها في النظام للتعبير عن المحتوى الموضوعي للوثائق، وتلك التي يتوقع أن يستخدمها المستفيد للتعبير عن حاجته إلى المعلومات.

أن الطريقة التي يتم فيها ترجمة مفاهيم المستفيد الموضوعية إلى مصطلحات بحثية محددة لصياغة طلب البحث، قد تعتمد على أدوات ووسائل مقننة أو تترك له حرية التعبير عن مفاهيمه البحثية بمفرداته الخاصة. ومن المميزات المهمة لنظم استرجاع المعلومات استجابتها للبحث باللغة الحرة والمقيدة في آن واحد. ويبدو إن هناك اختلاف في التعبير عن لغة الاسترجاع بين المتخصصين في مجال المعلومات حيث تستخدم مفاهيم مثل اللغة الطبيعية مقابل اللغة المقيدة، أو اللغة المقيدة مقابل اللغة الحرة، أو اللغة الطبيعية مقابل اللغة الاصطناعية ...الخ.

وتجدر الإشارة إلى أن اللغة الطبيعية سواء أكانت حرة (غير مقيدة) أم مقيدة هي ما يهمنا هنا .

اللغة الطبيعية الحرة:

وهى اللغة التي تجدها في الوثائق دون أي تعديل عليها، حيث عرفها لانكستر بأنها" اللغة الطبيعية او النص الحر أو الكلمات المفتاحية التي لا تنطوي على أية قيود لغوية أو مفردات معينة أو مصطلحات تفرض على المكشف أثناء عملية التكشيف أو المستفيد أثناء عملية البحث" حيث يكون للمكشف حرية الاختيار للكلمات الاهم مثل اختيار الكلمات المفتاحية الواردة في العنوان أو المستخلصات.

خصائص البحث باللغة الطبيعية الحرة

يرى البعض أن اللغة الطبيعية الحرة Natural Language Free لا تنطوي على أية قيود لغوية أو مصطلحات تفرض على المكشف أثناء عملية التكشيف أو المستفيد أثناء عملية البحث، وعليه فان نظام اللغة الطبيعية أو النصوص المطلقة لا يتم فيه أي نوع من أنواع التحكم في المصطلحات المستعملة. وقد تكون اللغة الطبيعية من وجهة نظر أخرى مجموعة المصطلحات المختارة من النص الحر والتي لا تخضع إلى أي تقييد. وهذه الحرية ليست مطلقة فهناك تقييد بضرورة وجود المصطلح البحثي في المحتوى النصي للوثيقة حتى يتم استرجاعها.

مزايا البحث باللغة الطبيعية الحرة :

توفر أوجه متعددة للبحث وبدرجة عالية من الشمول والتخصص ، فكل المفردات التي استعملها المؤلف يمكن أن تتحول إلى مصطلحات كشفية ، سواء أكانت أسماء أشخاص أو هيئات أو منظمات أو مواقع ...الخ.

قدرة النظام على استيعاب التطورات والتحديثات المستمرة للمفردات والمصطلحات الموضوعية، بسبب دخولها المباشر من نص الوثيقة الذي يحرص المؤلفون من خلاله مواكبة المستجدات العلمية والموضوعية .

بناء الكشاف من مفردات التسجيلات المقروءة آلياً اقل كلفة وأسهل في عملية الأعداد والتحديث .

يقدم مرونة أكبر في إجراء البحث بالربط اللاحق بين المفردات دون أن يدمر الهوية المستقلة لكل مفردة .

هذا النظام هو الأنسب في عمليات البحث التي يقوم بها المستفيد بشكل مباشر أو من خلال وسيط متخصص بالموضوع .

عيوب البحث باللغة الطبيعية الحرة :

في جانب آخر فان اللغة الطبيعية الحرة وكونها هي لغة الاتصال بين المؤلف بوصفه منتج المعلومات والمستفيد بوصفه المتلقي تتعرض إلى انتقادات كثيرة لدورها في استرجاع المعلومات، ويمكن تحديد أهم المشكلات التي تؤثر سلبا على كفاءة العمليات البحثية والمرتبط باستخدام هذه اللغة في صياغة استفسارات وطلبات البحث بالآتي :

تعدد المعاني : ويقصد به أن يكون للكلمة الواحدة أكثر من معنى، وهذه المشكلة تكاد تكون موجودة في معظم اللغات العالمية .

وجود المرادفات : تحدث عندما يتم التعبير عن الفكرة أو الشيء نفسه بكلمات أو عبارات مختلفة لفظا وكتابة ، وهذه المشكلة أوضح ما تكون في اللغة العربية* . مثال على ذلك ترجمة مصطلح CD-ROM إلى القرص المدمج، والمضغوط، و الليزري، والضوئي، والمكتنز.

الاستخدام المزدوج : بعض المفردات قد تستخدم في مجالات متعددة بحيث تكتسب مفهومها ودلالتها الموضوعية من المجال الذي تستخدم فيه بطريقة قد يختلف مفهومها كليا أو جزئيا من مجال إلى آخر . مثال على ذلك كلمة فايروس Virus بين الطب و الحواسيب.

من الصعوبة في هذا النظام تحديد العلاقات التي تربط فيما بين المصطلحات ودور كل منها.

مشكلة كتابة المصطلح الواحد بحالات مختلفة بين المفرد التثنية والجمع ، مما يؤدي إلى وجود لواحق حرفية لكل مصطلح وحسب نوع الحالة التي هو عليها.

إن نظام اللغة الطبيعية الحرة لا يمكن أن يعمل بكفاءة مقبولة دون وجود قائمة الكلمات الموقوفة(المستبعدة) Word List Stop التي تتطلب جهود كبيرة لغرض إعدادها في المستوى المقبول، مع التأكيد هنا إلى أن المفردات وكما أشرنا سابقا تكتسب مفهومها من المجال الذي تستخدم فيه، بحيث لا يكون لمفردة ما أهمية موضوعية في مجال معين و قد يكون لها أهمية في مجال آخر .

وبذلك تكمن مشكلات الاسترجاع باللغة الطبيعية الحرة أن نظام اللغة الحرة يوفر مرونة عالية في إجراء عمليات البحث واستجواب نظم استرجاع المعلومات، خاصة بالنسبة للمستفيد الذي يتمكن باستخدام هذه اللغة من استجواب النظام بواسطة استفسار في شكل جملة أو شبه جملة إذا لم يمتلك الخبرة في توظيف عناصر الربط البولياني، كما أن هذه اللغة في الواقع هي لغة الوثيقة أو لغة المؤلف لذا فان أمر استخدامها من قبل غير المتخصصين سهل، لكونهم لا يحتاجون إلى صياغة الألفاظ وإرجاعها إلى أصولها. على الرغم من هذا فان نتائج

العمليات البحثية التي نحصل عليها من خلال استخدام اللغة الحرة في صياغة طلب البحث لا تكون بمستوى الكفاءة المطلوبة.

وبشكل عام يمكن القول أن أهم المشكلات التي ترافق العمليات البحثية باستخدام اللغة الطبيعية الحرة (غير المقيدة) هي:

قد يفشل المستفيد في الحصول على نتائج مرضية بسبب افتقار اللغة إلى التحديد الدقيق للمصطلحات البحثية : " الأسلحة الجرثومية" أو "الأسلحة البايولوجية"

حالات فشل ناتجة عن وجود علاقات غامضة أو زائفة بين المفردات : "سرطان البحر".

حالات فشل ناتجة عن عجز اللغة من فك الارتباط ما بين المفاهيم والمفردات المتقاربة في المعنى : "المكتبات الخاصة" أو "المكتبات المتخصصة"

حالات فشل ناتجة عن عدم وجود بنيان مناسب مما يؤدي إلى تجاهل استخدام بعض المصطلحات التي تدعو الحاجة إليها لأغراض البحث الشامل: "فن العمارة" أو "الهندسة المعمارية".

البحث باستخدام هذه اللغة يتطلب التوسع في استخدام العوامل البوليانية، والتي قد يصاحبها حالات فشل بسبب جهل المستفيد بها أو بآلية استخدامها بالطريقة التي تحقق نتائج بحث مرضية.

حالات فشل ناتجة عن اختلاف محتمل في تهجئة المصطلح باللغة نفسها من دولة إلى أخرى Aluminum or Aluminium : .

اللغة الطبيعية المقيدة:

وهى اللغة المحكمة التي يتم تعيينها لمحتويات الوثائق خلال التحليل والاسترجاع، فهي لغة التكشيف المقيدة أو المقننة أو الاصطناعية تتضمن فهم المكشف لمضمون الوثيقة وتحيله إلى مفاهيم يعبر عنها بمصطلحات وتشمل(رؤوس موضوعات أو واصفات أو رموز).

خصائص البحث باللغة الطبيعية المقيدة:

عندما تحدد مفردات اللغة وتحسم مشكلاتها الدلالية وتستقر قواعدها النحوية من البداية وتضبط في قائمة معينة ويلزم الباحث والمكشف بالالتزام بها تصبح هذه اللغة مقيدة أو منضبطة Controlled Language.ويمكن أن تنظم هذه المفردات على شكل هيكل يبرز العلاقات الموضوعية واللغوية والنحوية بينها ويحدد الاستخدام الأمثل لها، وعلى أساس هذا التنظيم يمكن إعداد أدوات متنوعة مثل المكانز وقوائم رؤوس الموضوعات التي تشترك في كونها أدوات ضبط أو تقييد للغة الاسترجاع.

مزايا البحث باللغة الطبيعية المقيدة:

يمكن تلخيص مميزات ضبط المفردات في نظم استرجاع المعلومات على النحو الآتي:

تقدم مساعدة كبيرة للمستفيد أو الوسيط في اختيار المصطلحات الأكثر استخداماً للتعبير عن موضوع معين بما توفره من بنيان هرمي للمصطلحات واستعمال الإحالات

تحد إلى أدنى حد ممكن من تردد المستفيد في استخدام المصطلحات المناسبة، فهي تقدم مجموعة منتقاة من المصطلحات التي تم اختيارها من قبل متخصصين في مجال موضوعي محدد للتعبير عن محتوى الانتاج الفكري في أحد المجالات

توفر الأدوات المساعدة بذلك يكون البحث باللغة المقيدة أسهل ويتطلب قدراً ضئيلا من الجهد المبذول من قبل المستفيد والوسيط

تحد هذه اللغة من التباين أو الاختلاف في وجهات النظر في استخدام المصطلحات المناسبة بين المستفيد والوسيط .

تقلل من الحاجة إلى استخدام العوامل البوليانية في ربط المصطلحات والمفاهيم البحثية

عموما البحث باللغة المقيدة اقل كلفة عندما يكون الوقت هو العامل الذي تحتسب عليه في إجراء العمليات البحثية.

عيوب البحث باللغة الطبيعية المقيدة :

على الرغم مما للّغة المقيدة من ميزات وكما أشرنا إليها سابقا إلا أن معظم نظم استرجاع المعلومات تتطور اليوم باتجاه تيسير سبل البحث باللغة الطبيعية الحرة ، ويرجع السبب في ذلك إلى جملة من

العوامل التي يراها الباحث على إنها عيوب مباشرة وغير مباشرة للغة المقيد، وكما يأتي :

أن أي محاولة تفرض على المستفيد في استخدام اللغة المقيدة في صياغة استفساره قد تؤدي إلى استقراره على اقل مما يريده فعلاً

التغطية الشاملة للنتائج التي توفرها اللغة المقيدة، غالبا ما تكون على حساب مستوى التخصص، بعبارة أخرى فان الاستدعاء العالي يكون دائما على حساب الدقة .

التكاليف العالية لبناء وتطوير أدوات ضبط المصطلحات من مكانز وقوائم رؤوس موضوعات. وبسبب هذه التكاليف يصعب في كثير من الأحيان تحديث هذه الأدوات لتكون مواكبة للتطورات الجارية في المجالات العلمية المختلفة، مما يؤدي إلى وجود ضعف في تغطيتها للمصطلحات المستحدثة

نظم البحث باللغة المقيدة تحد من إمكانية المستفيد في إجراء العملية البحثية دون وساطة، وتدخل الوسيط غالبا ما يترتب عليه زيادة في تكلفة تنفيذ هذه العمليات.

أن أي خطا يرتكبه المكشف في اختيار المصطلح الموضوعي المعبر عن المحتوى الموضوعي للوثائق، يؤدي إلى خسارة محتملة عند الاسترجاع لتسجيلات يمكن أن تكون ذات صلة بموضوع البحث.

ج- اللغة الهجينة:

وهى مزيج من المصطلحات المنضبطة التي خصصها المكشف واللغة الطبيعية وهى العبارات والكلمات الواردة في العناوين أو المستخلصات والتي هي لغة المؤلف أو لغة الوثيقة عادة"

ويعرفها لانكستر بأنها" اللغة التي تجمع بين اللغة المقيدة واللغة الطبيعية وغالبا تمثل اللغات المقيدة العريضة نسبيا ويتم تكشيف النظام بواحدة أو أكثر من الواصفات العريضة ومصطلحات اللغة الطبيعية المستخلصة من عناوين الوثائق أو من نصوصها أو من كليهما".

والاتجاه السائد الآن بوجه عام يسير إلى توفير مقومات البحث باللغة الطبيعية واللغات المقيدة في الوقت نفسه في نظم استرجاع المعلومات الآلية ، ولتوفير مقومات الطريقتين ما يبرره في الطبيعة المختلفة لهذه اللغات، وهناك العديد من الدراسات التي تؤكد على الطابع التكاملي لكل من اللغة الحرة واللغة المقيدة في الاسترجاع ، فمن الممكن دائما أن تكفل كل لغة من هذه اللغات استرجاع مجموعة من التسجيلات ذات الصلة بالموضوع لم يكن وبالإمكان الوصول إليها باستخدام اللغة الأخرى وهذا هو المقصود بالتكامل.

12/ الجوانب المؤثرة في كفاءة نظم استرجاع المعلومات:

هناك جوانب كثيرة مرتبطة بمدى دقة نظام استرجاع المعلومات على تلبية الاحتياجات المعلوماتية للمستفيد بشكل ملائم، وتتنوع تلك

الجوانب ما بين التغطية الموضوعية، التكشيف، إستراتيجية البحث، على النحو التالي:

❖ **مدى التغطية الموضوعية:**

المقصود هنا مدى توفر المصادر أو التسجيلات الببليوجرافية التي تلبي الاحتياجات المعلوماتية للمستفيد ويعتمد ذلك على عدد المصادر ونوعيتها، ومدى اكتمالها بالنسبة لاحتياجات المستفيد.

❖ **إجراءات التكشيف:**

يتيح التكشيف الجيد استرجاع المعلومات التي تلبي احتياجات المستفيد، حيث يرتبط التكشيف بعوامل كثيرة مثل

1) مدى التخصيص في اللغة: بمعنى ما إذا كانت اللغة المستخدمة في التكشيف تميل إلى التعميم أم التخصيص ؟
2) مستوى الربط في اللغة : أي مدى الربط المسبق للكلمات المفردة في صيغ أكثر تعقدا عن طريق استعمال العبارات أو الرؤوس الفرعية أو الأدوات اللغوية الصريحة مثل الروابط.
3) مقدار ما يقدم صراحة للمكشفمن تنظيم للعلاقات؛ في شكل علاقات دلالية بين المصطلحات مثل المصطلحات الأعرض أو الأضيق أو المفصلة.

وتتعدد أسباب أخطاء عملية التكشيف فمنها:

✓ فشل المكشف فى تحليل المفهوم بطريقة صحيحة وعدم فهمه موضوعها بدقة.

- ✓ فشل المكشف فى ترجمة المفاهيم واختياره مصطلحات غير صحيحة للتعبير عن الموضوعات.
- ✓ حذف المكشف لجوانب مهمة من المفاهيم التى تنطوي عليها الوثيقة.
- ✓ استخدام المكشف مصطلحات عامة للتعبير عن الوثيقة بسبب الافتقار إلى المصطلحات المتخصصة الملائمة.
- ✓ استخدام المكشف مصطلحات عامة للتعبير عن المفاهيم على الرغم من توافر المصطلحات المتخصصة.
- ✓ كما أن الأخطاء الإملائية أو النحوية التي توجد بقواعد البيانات تكون ناجمة عن الإدخال غير الصحيح من جانب مدخل البيانات، والعاملين على لوحة المفاتيح الذين ينتجون النص المقروء آليا. فجميعهم عرضة للخطأ، والأخطاء سوف تؤثر على التحقيق من حيث أننا سوف نسترجع إجابات خاطئة، كما أنها سوف تؤثر على الاستدعاء حيث يتم فقد بعض الوثائق كان يجب أن يتم استرجاعها.

❖ **إستراتيجية البحث:**

تؤثر إستراتيجية البحث المطبقة لاسترجاع المعلومات على مدى الدقة في الحصول على النتائج المرغوبة، وسواء كان اختصاصي المعلومات أو المستفيد يقوم بعملية البحث؛ فإن هناك بعض الجوانب التي تؤثر على دقة الاستدعاء.

1) مدى الخبرة والإبداع والكفاءة المتوافرة لدى القائم بعملية البحث.
2) نوع المصطلحات المستخدمة في كشاف النظام؛ إذا كانت المصطلحات المقيدة هي المستخدمة فإن المستفيد–اختصاصي المعلومات لا يستطيع أن يخرج عن هذا الإطار.
3) عدم تحديد المستفيد للحاجات المعلوماتية الحقيقية خلال مقابلة ما قبل البحث.
4) عدم جودة بناء استراتيجية البحث.
5) الإخفاق فى إجراء المقابلة المرجعية من قبل المسئول.
6) البحث تحت مصطلحات خاطئة.
7) اختيار قاعدة بيانات غير ملائمة.
8) عدم استخدام مترادفات/ مشتقات/متصرفات مصطلح البحث.

الفصل الثاني

تقييم نظم استرجاع المعلومات

مقدمة:

إن تصميم وتطوير نظام لاسترجاع المعلومات، يرتبط أساساً بتحقيق هدف أو مجموعة أهداف محددة، والتي ينظر إليها على إنها غاية النظام ، وفي البدايات الأولى لنظم استرجاع المعلومات المحوسبة كانت قدرة النظام على استرجاع المعلومات هي هدفه الأول وغايته الأهم. ولكن بعد مرور مدة قصيرة من الزمن تبين أن استرجاع المعلومات كعملية مجردة وغير موجهه لا يمكن أن تكون هدفاً أوحداً للنظام ، طالما إنها لا ترتبط بحاجة موضوعية أو بحثية لمستفيد ما ، لذا فان حكم المستفيد بصلاحية المعلومات المسترجعة من النظام استجابة لمطلبه البحثي أصبحت في مقدمة أهداف نظم استرجاع المعلومات، وترتبط بها مقاييس تقييم نظم استرجاع المعلومات التقليدية والآلية على حد سواء، وعلى أساسها تتم المفاضلة بين مختلف أنواع النظم .

وتجدر الإشارة هنا إلى إن قدرة نظم استرجاع المعلومات على إمداد المستفيد بالمعلومات المناسبة من بين الكم الهائل من المعلومات المخزنة في وسائطها التخزينية غاية صعبة المنال ولا يمكن مقارنتها مع قدرة تلك النظم على استرجاع المعلومات، بغض النظر عن ملاءمتها أو عدم ملاءمتها . وعليه فقد أصبح من ضروريات عمليات التقييم إشراك المستفيد صاحب الحاجة الفعلية للمعلومات في الحكم

على صلاحية المعلومات المسترجعة من النظام ، فالمعلومات لا قيمة لها إذا لم ترتبط باهتمام أحد المستفيدين أو حاجته الموضوعية

ولصعوبة قياس دوافع المستفيدين لطلب المعلومات بطرائق كمية، وكذلك التباين الكبير بين مستفيد وآخر في الحكم على صلاحية نفس المعلومات المسترجعة استجابة للموضوع ذاته، أضحت عملية تقييم نظم استرجاع المعلومات معقدة في إجراءاتها وغير دقيقة في نتائجها ، لهذا اجتهد العديد من الباحثين والمختصين في ابتكار معايير وأساليب مختلفة ، الهدف منها إجراء تقييم شامل لمختلف أنواع نظم استرجاع المعلومات ، وتوجيه العمليات البحثية بشكل متوازن بين نوع وكم المعلومات التي يحصل عليها المستفيد ومقدار الجهد والوقت المبذولين لغرض الحصول على تلك المعلومات ، فضلا عن الكلفة المادية التي قد يتحملها المستفيد بشكل مباشر أو غير مباشر من خلال الرسوم المترتبة على مؤسسات المعلومات مقابل اشتراكها في خدمات نظم استرجاع المعلومات

ومنذ عام 1950 ظهرت في أدبيات الموضوع العديد من الدراسات التي تعالج مشكلة استرجاع المعلومات، وتقييم نظم المعلومات، ومن ابرز هذه المحاولات دراسات كرانفيلد عام 1957 التي كانت تهدف إلى وضع منهاج لقياس الكفاءة النسبية لمختلف لغات التكشيف وقياس التأثير النسبي للمداخل الكشفية باستخدام أساليب الاستدعاء والدقة ومع

التطور الكبير الذي حصل في صناعة نظم استرجاع المعلومات الآلية واتساع نطاق استخدامها وكذلك تنوع شرائح المستفيدين منها ، ظهرت العديد من أساليب ومقاييس تقييم نظم الاسترجاع والتي تهدف في مجموعها إلى قياس الكفاءة النسبية لهذه النظم.

1/ مستويات تقييم النظم:

سبقت الإشارة إلى إن نظم استرجاع المعلومات المحوسبة تطورت بجهود المؤسسات التجارية التي وضعت الربح المادي في مقدمة أهدافها، والتي كانت تعمل بشكل متواصل إلى إيجاد مبررات لاستمرار خدماتها وزيادة عدد مستخدميها، ومن المؤكد إن استمرار أي نظام لاسترجاع المعلومات يرتبط بقدرته على تلبية احتياجات المستفيدين من المعلومات الملاءمة لمتطلباتهم البحثية والموضوعية، من جانب أخر فان الوصول إلى المعلومات بمدة زمنية قصيرة وبجهد أقل يبذله المستفيد سواء كان صاحب الحاجة الفعلية للمعلومات أم وسيط العملية البحثية، وبكلفة اقتصادية أقل، تعد عوامل مضافة إلى صلاحية المعلومات المسترجعة عند إجراء عملية التقييم . وبشكل عام هناك ثلاث مستويات محتملة يمكن من خلالها وعلى أساسها تقييم نظم استرجاع المعلومات وهي :

أ- المستوى الأول : تقييم فاعلية النظام التي تتصل باعتبارات رضا المستفيد.

ب- المستوى الثاني: تقييم فاعلية التكلفة التي تتصل برضا المستفيد عن الكفاءة الداخلية للنظام .

ت- المستوى الثالث : تقييم عائد التكلفة المتصلة بأهمية النظام مقابل تكاليف تشغيله

ويمكن القول أن مستويات التقييم تتصل عموما بمقاييس التكلفة، سواء تم حساب تلك التكلفة من جانبها المالي أو مقدار الجهد المبذول عند إجراء البحث، أو كلاهما، مع حساب الوقت المستنفذ في إجراء العملية البحثية من لحظة تقديم الطلب إلى حصول المستفيد على نتائج البحث. وطالما كان المستفيد هو الطرف الذي يتحمل تكلفة البحث، فإن اعتبارات الجودة سيكون لها دور مهم في مستويات تقييم النظم خاصة تلك التي تتصل بحداثة المعلومات المتوفرة في النظام والتغطية الموضوعية والمصدرية له ، مع التركيز على تكامل المعلومات ودقتها.

أما إذا نَظر إلى مقاييس التكلفة من وجهة نظر الإدارة التشغيلية فسنجد إنها تعتمد على الموازنة بين عائدات النظام التي تمثل (إجمالي الإيرادات المالية المحصلة من تشغيل النظام خلال مدة زمنية معينة) مقارنة بتكاليف تشغيله (إجمالي المصروفات التي تتحملها الإدارة التشغيلية أو المؤسسة التمويلية في سبيل تشغيل النظام خلال المدة نفسها) ، و معدل الفرق المقبول بين العائد والتكاليف يرتبط بسياسة المؤسسة التمويلية فيما إذا كانت مؤسسة ربحية أم غير ربحية، وفي كل

الأحوال فان استمرار عمل النظام مرتبط ببقاء معدل التكاليف دون معدل العائد من النظام إلا إذا كان هناك دعم من نوع ما لاعتبارات خاصة بالمؤسسة أو بنوعية المستفيدين من خدمات النظام . وعلى هذا الأساس يمكن النظر إلى مستويات التقييم بكونها مقاييس، ترتبط بما يأتي :

- رضا المستفيد عن نتائج العملية البحثية (المخرجات) التي يحصل عليها والتي تتم مقارنتها بمعدل الجهد المبذول والوقت المستنفذ، والكلفة المادية المدفوعة مقابل الخدمة.
- رضا مشغلي النظام من وسطاء ومنفذين عن إجمالي ما يحصلون عليه من مبالغ نظير خدماتهم مقارنة بحجم الجهد المبذول في المراحل المختلفة للعمليات البحثية.
- رضا المؤسسة التمويلية عن العائدات مقارنة مع تكاليف تشغيل النظام.

2/ العوامل والمقاييس النوعية:

في سبيل الوصول إلى تقييم موضوعي ودقيق لكفاءة نظم استرجاع المعلومات ، لابد من أن يكون للمستفيد دور مباشر فيه من خلال حكمه على نوعية المخرجات التي يحصل عليها بوصفها نتائج للعملية البحثية التي يقوم بها، وعلى هذا الأساس يرى لانكستر إن الإجابة على

التساؤلات الآتية التي ترتبط بالمستفيد تشكل مقياساً نوعياً عن مدى كفاءة هذه النظم:

أ- هل يحصل المستفيد على ما يبحث عنه فعلاً أم لا؟

ب- إلى أي حد يحصل عليه مكتملا أو دقيقاً؟

وإذا كان التساؤل الأول يتصل بقدرة النظام على الاسترجاع (الاستدعاء) ، فإن التساؤل الثاني يتصل بقدرة النظام على استرجاع المعلومات المناسبة فقط واستبعاد ما عداها (الدقة) وفي كلا الحالتين يجب أن يكون هناك تقييم لمقدار الجهد الذي ينبغي أن يبذله الوسيط أو المستفيد للحصول على المعلومات المناسبة، مع حساب مقدار الوقت المستنفذ ما بين صياغة الاستفسار وتلقي النتائج المناسبة.

وفي منتصف خمسينيات القرن الماضي تم وضع عدد من العوامل لتستخدم مقاييس نوعية عند إجراء عملية التقييم وهي:

- عامل الاتفاق(Resolution) ويقصد به مجموعة التسجيلات التي يوجه النظام بهدف استرجاعها

- عامل العزل(Elimination) ويقصد به مجموعة التسجيلات التي يوجه النظام بعدم استرجاعها.
- عامل الملاءمة (Pertinence) ويقصد به مجموعة التسجيلات التي استرجعها النظام وتبين بعد الاختبار ملاءمتها .
- عامل التشويش(Noise) ويقصد به مجموعة من التسجيلات التي استرجعها النظام و تبين بعد الاختبار عدم ملاءمتها .
- عامل الاسترجاع(Retrieval) ويقصد به مجموع التسجيلات التي قام النظام باسترجاعها فعلا من مجموع التسجيلات الموجودة في النظام
- عامل الحذف(Omission) يقصد به مجموعة التسجيلات المناسبة في النظام والتي عجز النظام عن استرجاعها.

وهذه المقاييس يمكن تعريفها على أساس أربع مقادير كمية هي:

- عدد التسجيلات المخزنة في النظام، ويرمز لها بالحرف N.
- عدد التسجيلات التي يسترجعها النظام استجابة لاستفسار ما وتحظى بأهمية مناسبة، ويرمز لها بالحرف M.
- عدد التسجيلات التي كانت أهمية فعلية مناسبة بعد الاختبار الشخصي للمجموعة(M)، ويرمز لها بالحرف W.

- عدد التسجيلات ذات الأهمية الفعلية المناسبة التي تدخل ضمن مجموعة التسجيلات (N)، ويرمز لها بالحرف X.

وعلى أساس هذه المقادير يمكن حساب العوامل السابقة باستخدم المعادلات الرياضية وكما يأتي:

- عامل الاتفاق=N/M
- عامل العزل=N–M/N
- عامل الملاءمة=W/M
- عامل التشويش=M–W/M
- عامل الاسترجاع=W/X
- عامل الحذف=X–W/X

3/ مقاييس الاستدعاء والدقة:

ما تزال مقاييس الاستدعاء Recall و الدقة Precision تحظى بقبول واسع في مجال التقييم النوعي لنظم استرجاع المعلومات ، وذلك بسبب ارتباطهما برضا المستفيد وطريقة حكمه على مخرجات العملية البحثية أولا ، و ترابطهما مع بعض ثانياً، حيث يعد معدل الاستدعاء مقياس لقدرة النظام على استرجاع التسجيلات المتصلة بموضوع البحث

، بينما يعد معدل الدِقة مقياس لقدرة النظام على استبعاد التسجيلات التي ليست ذات صلة بموضوع البحث.

وعلى أساس ما تقدم فإن الاستدعاء بوصفه مقياس نوعي يمكن أن يعبر عن قدرة النظام في استرجاع المعلومات المناسبة حول موضوع البحث ممثلة بعدد التسجيلات المسترجعة والتي تعكس الوثائق المخزنة في مستودع النظام من المجموع الكلي للوثائق حول ذات الموضوع، أي هو نسبة الإصابات إلى مجموع المواد المناسبة في النظام، وبالإمكان التعبير كميا عن درجة نجاح النظام في استرجاع التسجيلات ذات الصلة بموضوع البحث بواسطة نسبة الاستدعاء، وتعرف نسبة الاستدعاء على أنها عدد التسجيلات ذات الصلة بالموضوع التي استردجعها النظام على إجمالي عدد التسجيلات ذات الصلة بالموضوع والموجودة في النظام مضروبة في (100)، أما نسبة الدقة فهي عدد التسجيلات ذات الصلة التي استرجعها النظام على إجمالي عدد التسجيلات التي استرجعها النظام مضروبة في (100).

❖ **معيار الاسترجاع(الاستدعاء):**

المقصود بالاسترجاع هنا قدرة النظام على استرجاع المعلومات المناسبة حول موضوع معين ممثلة بعدد التسجيلات التى تعكس الوثائق المخزونة فى مستودع المعلومات الآلى من مجموع الوثائق الكلى حول ذات الموضوع، أى هو نسبة الإصابات إلى مجموع المواد المناسبة، والتى يمكن احتسابها بالطريقة التالية

$$\text{معدل الاستدعاء} = \frac{\text{عدد الوثائق المناسبة المسترجعة}}{\text{مجموع الوثائق المناسبة في النظام}} \times 100$$

فعلى سبيل المثال إذا تم استرجاع ما مجموعه(980) تسجيلة استجابة لبحث عن موضوع الوصول الحر للمعلومات وكان إجمالي عدد التسجيلات عن هذا الموضوع فى مستودعات النظام (1000) تسجيلة، فإن احتساب معدل الاسترجاع يكون:

980

المعدل = _____ x100= 98%

1000

وفقا لهذا المعيار فإن النظام المثالي و النظام القادر على تحقيق نسبة استرجاع 100% أي استرجاع كل التسجيلات الموجودة فى مستودع خزن النظام عن موضوع معين استجابة لبحث عن ذات الموضوع، وإن أي فشل في تحقيق هذه النسبة يعنى مؤشرا سلبيا على النظام.

إلا أننا نشير هنا إلى أن نظم استرجاع المعلومات الآلية تحتوى في مستودعاتها على كمية كبيرة من التسجيلات التي تعكس الموضوعات المختلفة والتى تم انتخابها لتكون قاعدة بيانات بحثية تستخدم من قبل المستفيدين .لكن ليس من السهل معرفة العدد الكلى للتسجيلات عن موضوع ما، وهل تم استرجاع كافة ما موجود فى القاعدة أم لا، لذا يصعب تقييم النظم الآلية خاصة من قبل المستفيدين وذلك لعدم وجود وسيلة مناسبة تمكننا من معرفة التفاصيل الكلية عن

حجم المدخلات عن كل موضوع على حدة، وهناك مجموعة من العوامل التي تؤدي إلى إخفاق في عمية الاسترجاع هي:

- التكشيف: مثل تجاهل المكشف لبعض الجوانب الموضوعية المهمة واختيار مصطلح غير مناسب، أو أن يكون العنوان فقط هو مصدر المكشف.
- لغة التكشيف: حيث قد يكون هناك قصور فى بيان العلاقة بين المصطلحات فى لغة التكشيف.
- استراتيجيات البحث: فقد يخطئ المستفيد فى بيان العلاقات الموضوعية بين المصطلحات البحثية وعد الدقة فى اختيار العناصر المناسبة للمنطق البوليانى للربط بين المصطلحات.
- النصوص المدخلة: حيث أن الأخطاء الإملائية واللغوية وترادف المفاهيم من العوامل المؤثرة على إخفاق عملية الاسترجاع.

❖ **معيار الدقة:**

أما فى ما يخص معيار الدقة فيقصد به نسبة عدد التسجيلات المناسبة التى تم استرجاعها من مجموع التسجيلات المسترجعة، ويمكن الحصول على الدقة من خلال العملية الرياضية التالية:

عدد التسجيلات المناسبة المسترجعة

معدل الدقة=_____x100

العدد الكلي للتسجيلات المسترجعة

فنفرض أن عدد التسجيلات المناسبة من مجموع ما تم استرجاعه عن موضوع الوصول الحر للمعلومات هو (98) تسجيلة، وعليه يمكن احتساب معدل الدقة كالآتي:

98

معدل الدقة=_____x100

980

إن قياس مستوى الدقة أسهل من الاسترجاع، إذا أن التعرف على التسجيلات المناسبة يتم من خلال الاطلاع الفعلي على مجموعة من التسجيلات المسترجعة فعلا، ولا توجد دقة إذا لم يكن هناك استرجاع،

لكن الحكم على دقة التسجيلات وصلاحيتها لموضوع البحث يعتمد أساسا على ردود أفعال المستفيدين التي في الغالب لا يمكن الوثوق بها تماما خاصة في قواعد البيانات التي توفر مخرجات على شكل معلومات ببليوجرافية أو مستخلصات، والتي يصب الحكم على صلاحيتها.

لهذا فإن مقاييس الاستدعاء والدقة قد لا تعطي نتائج دقيقة عند تقييم نظم الاسترجاع المعلومات خاصة إذا كان المستفيد هو صاحب القرار الوحيد في الحكم على صلاحية أو عدم صلاحية المعلومات المسترجعة، من ناحية أخرى فإن نسبة الاستدعاء قد لا يمكن قياسها بصورة دقيقة، لصعوبة تحديد عدد التسجيلات ذات الصلة بموضوع البحث الموجودة في النظام ولم يتم استرجاعها .

4 / مفاهيم الكفاءة:

لاشك أن مفهوم الكفاءة هنا لا يرتبط بقدرة النظام على الاسترجاع فحسب ، وإنما في قدرته على إشباع حاجات المستفيدين من المعلومات ذات الصلة بموضوعاتهم البحثية أو ما يعرف اصطلاحا "اكتمال

البحث" ، فليس من أهداف عمليات تقييم نظم استرجاع المعلومات قياس قدرتها على الاسترجاع باعتبارها عملية نهائية ، لان قدرة النظام على الاسترجاع تعد وسيلة المستفيد لاختيار التسجيلات التي يراها وثيقة الصلة بالموضوع ، ولا يمكن النظر إليها على إنها غاية النظام . فمفهوم كفاءة الاسترجاع هو مصطلح له دلالة أوسع ومعنى اشمل مما يبدو عليه، ولعل السبب في ارتباط الكفاءة بمفهوم الاسترجاع جاء كنتيجة طبيعية لارتباط مقاييس التقييم التي أشرنا إليها سابقاً بمخرجات النظم التي يتم استرجاعها استجابة لمطلب بحثي محدد .

لكن هذا لا يعني إن كفاءة الاسترجاع عملية منفصلة عن كفاءة النظام ككل، فكل جزء من أجراء نظم استرجاع المعلومات سواء كان مادياً أم برمجياً أم تشغيلياً يؤثر ويتأثر بأجزاء النظام الأخرى، ولضمان الدقة في التعبير لابد من الفصل في المعنى بين مفاهيم "كفاءة البحث" و "كفاءة الاسترجاع" و "كفاءة النظام" وكالآتي :

أ/كفاءة البحث : ترتبط بنتائج العملية البحثية التي يقوم بها المستفيد ويتم تقييمها وقياسها بشكل مستقل عن نتائج العمليات البحثية لمستفيدين آخرين في زمن سابق أو لاحق له، وسيكون المستفيد صاحب الحاجة الفعلية للمعلومات هو الحَكَم الفصل على كفاءة البحث

، والعامل المؤثر الأول فيها . وبشكل عام يمكن أن يسترجع المستفيد مجموعة من التسجيلات تضاهي إستراتيجية البحث التي تم بواسطتها استجواب النظام والتي يمكن تصنيفها مقارنة بموضوع البحث واعتمادا على رأي المستفيد إلى:

- ✓ مجموعة من التسجيلات ذات الصلة الوثيقة بموضوع البحث (تطابق).
- ✓ مجموعة من التسجيلات ذات الصلة القريبة بموضوع البحث (تقارب).
- ✓ مجموعة من التسجيلات ذات الصلة البعيدة بموضوع البحث (تباعد)
- ✓ مجموعة من التسجيلات ليست ذات صلة بموضوع البحث (ضوضاء)

ب-كفاءة الاسترجاع : تقييم أشمل يحتاج إلى دراسة نتائج مجموعة من العمليات البحثية لعدد من المستفيدين وباستخدام المفاهيم البحثية نفسها ، ثم مقارنة النتائج للوصول إلى قرار بخصوص كفاءة أو عدم كفاءة الاسترجاع ، وبذلك فإن جوهر الاختلاف عن كفاءة البحث تكمن في تَحديد الرغبات والدوافع غير المعلنة للمستفيد في الحكم على

صلاحية النتائج. ويعتبر دور الوسيط من العوامل المؤثرة في كفاءة الاسترجاع إلى جانب دور المستفيد بالنسبة للعمليات البحثية التي تجري بالوساطة، ويتم قياس كفاءة الاسترجاع عادة باستخدام مقاييس " الاستدعاء والدقة "، والواقع إن معدل الدقة هو المقياس المستخدم في قياس كفاءة الاسترجاع، فمن غير الممكن التحقق من معدل الاستدعاء في هذه المرحلة لصعوبة تحديد عدد التسجيلات ذات الصلة بالموضوع والتي لم يتم استرجاعها. وبشكل عام تزداد كفاءة الاسترجاع كلما اقترب الفرق بين العدد الكلي للتسجيلات التي تم استرجاعها وبين عدد التسجيلات التي حكم المستفيدون بصلاحيتها للموضوع من الصفر.

ج- **كفاءة النظام** : من الصعب أحيانا الفصل بين كفاءة النظام وكفاءة الاسترجاع إلا في حدود قدرة النظام على تحقيق نتائج متطابقة لعمليات بحث مختلفة بغض النظر عن قدرة المستفيد أو الوسيط في إدارتها بالطريقة التي تؤدي إلى رفع كفاءة الاسترجاع ، وتشكل الكفاءة العلمية والعملية لنظام استرجاع المعلومات مقياساً لقابلية النظام للقيام بالوظائف التي صمم من أجلها على أساس إن من يستجوب النظام خبير في هذا المجال.

ففي الوقت الذي يتم فيه قياس كفاءة الاسترجاع بمقاييس نوعية ترتبط بمخرجات النظام ، فان كفاءة النظام تتأثر إلى جانب ذلك بعوامل المدخلات والتي من أهمها:

- ✓ نوع الوثائق التي يحتوي عليها النظام .
- ✓ إلى أي حد يمكن التعرف على المحتوى الموضوعي لهذه الوثائق.
- ✓ مدى صلاحية لغة النظام في التعبير عن المحتوى الموضوعي لهذه الوثائق.
- ✓ مرونة برمجيات البحث.
- ✓ سهولة التعامل مع واجهات المستفيد بالدرجة التي يقلل فيها النظام إلى أدنى حد ممكن تأثير خبرات المستفيد على نتائج العملية البحثية.

وعلى الرغم من صعوبة تحقيق موازنة نسبية بين مستويات الاستدعاء ومستويات الدقة بالتخلص من التناسب العكسي بينهما والوصول إلى إمكانية استرجاع كل التسجيلات ذات الصلة بموضوع البحث، الموجودة في النظام والتي يقرر المستفيد ملاءمتها للموضوع .

5/ العوامل المؤثرة في الكفاءة:

بغض النظر عن نوع ومفهوم الكفاءة فيما إذا كانت كفاءة بحث أم استرجاع أم نظام فأنها تشترك جميعاً بوجود عوامل ذات تأثير مباشر على ارتفاع أو انخفاض مستوى هذه الكفاءة ، وبدءا برغبة المستفيد في البحث عن المعلومات لإشباع حاجاته الموضوعية المرتبطة بنشاطه البحثي والعلمي والعملي ، ومروراً باستخدامه لنظم استرجاع المعلومات المتاحة له ، وانتهاءً بحصوله على المعلومات التي يقرر صلاحيتها لموضوع البحث ، تكون هناك سلسلة طويلة من الأنشطة والفعاليات والعمليات والوظائف التي يرتبط قسم منها بالمستفيد صاحب الحاجة الفعلية للمعلومات ، وقسم آخر بالوسيط الذي يمثل حلقة الوصل بين النظام والمستفيد ، وقسم ثالث بأجزاء نظام استرجاع المعلومات كافة ، والتي لاشك سيكون لها تأثير إيجابي أو سلبي على نتائج العمليات البحثية عموماً التي ستنعكس على كفاءة الاسترجاع .

وبغض النظر فيما إذا كان المستفيد هو نفسه من يقوم بأجراء عملية البحث أو يفوض أخصائي المعلومات بدلا عنه، فإن العوامل المشار لها تكاد تمثل الحالة العامة والشائعة للعمليات البحثية ، مع التأكيد على احتمال وجود تباين نسبي في العوامل المؤثرة في الكفاءة بين حالة وأخرى، وهو ما سيتم استعراضه في الصفحات اللاحقة من خلال دراسة

دور كل من المستفيد النهائي والوسيط واختيار قاعدة البيانات وتأثيراتها على كفاءة الاسترجاع.

6/ تأثير دور الوسيط على كفاءة الاسترجاع:

دور الوسيط أو أخصائي المعلومات كان له وجود حقيقي في ميدان العمل التقليدي لمؤسسات المعلومات، ولكن تحت مسميات مختلفة في أشكالها ومتماثلة في مضامينها، وطبيعة الدور المنوط بها، فلقد كان مسؤول الخدمة المرجعية يقدم وساطة عالية المستوى، فلم يكن دوره في تسهيل الوصول إلى المعلومات فقط، وهو عمل يستطيع المستفيد المطلع القيام به، و إنما يتعدى ذلك إلى استثمار كل خبرته في مجال مصادر المعلومات لتقديم إجابات منتقاة، والعمل على توفير المعلومات الملاءمة لحاجة المستفيد، وهو عمل أبعد في مداه من توفير مصادر هذه المعلومات .

وبعد التطور الكبير الذي حدث في مجال بث المعلومات وتداولها باستخدام تقنيات الحواسيب وظهور خدمات البحث بالاتصال المباشر والبحث في قواعد البيانات على الأقراص المدمجة واستخدام شبكة الإنترنت ، شهدت العلاقة بين أخصائي المعلومات والمستفيد تحدي كبير، بعد أن انتزعت منه سلطة امتلاك المصادر المرجعية لمصلحة المستفيد.

فسهولة استخدام الأجهزة الحديثة وبرامجها من خلال واجهات بحث بسيطة ومرنة وكذلك تنامي النشر الإلكتروني الذي يضع المعلومات مباشرة في خدمة المستفيد، كل ذلك يوحي باحتمال ضعف وساطة أخصائي المعلومات للدرجة التي قد تؤدي إلى الاستغناء عن دوره نهائيا. لكن من المهم التأكيد هنا على أن الكمية الهائلة للمعلومات المتاحة بواسطة التكنولوجيات الحديثة وكثافة الكلمات الدالة التي تسمح بها أدوات البحث الآلية، تؤكدان من جديد إلى أهمية دور أخصائي المعلومات في طرح الأسئلة البحثية وبناء استراتيجيات البحث المتقدمة. ويبقى المستفيدون بوجود هذه التقنيات بحاجة إلى من يساعدهم في الحصول على المعلومات بسرعة ودقة أكبر مما لو نفذوا العملية بأنفسهم، فواجهات النظم الصديقة، ومحركات البحث لم تكن وسائل كافية ليتمكن المستفيد من خلالها إنجاز العملية البحثية بالاعتماد المباشر على مهارته وخبراته في إدارة البحث، فهذه الأدوات قدمت أنصاف الحلول، واحتفظ أخصائي المعلومات بالنصف الآخر من الحل، المتمثل في قدرته على التعرف بوضوح أكبر على احتياجات المستفيدين والتعبير عنها بمفاهيم ومصطلحات دقيقة، تضمن للمستفيد فرصة الحصول على نتائج بحثية ملاءمة مع تقليص كبير في الوقت والجهد المستنفد بإجراء التحاور مع نظم استرجاع المعلومات.

أ/ الوساطة في مجال العمل التقليدي:

الوساطة في مجال المعلومات لها معنى قد يختلف من الناحية التطبيقية عن أنماط الوساطة في المجالات الأخرى، فالوسيط خارج حدود عمل مؤسسات المعلومات يقوم بدوره بهدف تحقيق مكسب مادي أو معنوي يحصل عليه بصورة مباشرة أو غير مباشرة . لكن مجمل المهام والوظائف التي تنجز من قبل العاملين في هذه المؤسسات لا تخرج عن كونها نمطاً من أنماط الوساطة. المكتبيون على سبيل المثال هم وسطاء بين مصادر المعلومات والمستفيدين، وتحقيق رضا المستفيد هو الهدف الأول للوسيط، ففي ظل أنماط النشر التقليدي للمعلومات كانت هناك حاجة ملحة إلى وسيط يعمل على اختيار وتنظيم المعلومات المنشورة في مختلف أنحاء العالم و إعدادها بالطريقة التي تحقق الاستفادة منها، ومع التطورات المتلاحقة في عمل مؤسسات المعلومات، اتسع نطاق الوساطة ليشمل عملية الاستخلاص و التكشيف والتحليل بهدف تسهيل عملية البحث على المستفيد، بعد أن أصبح الاتجاه السائد في أوساط المستفيدين يتجه نحو البحث عن المعلومات لا عن مصادرها، كل ذلك أدى إلى ظهور وظيفة أخصائي المعلومات التي تنطوي تحتها مسميات الوظائف والمهام الآتية:

أ- **معالج المعلومات** : وهو الشخص المسؤول عن رسم أطر المعلومات وتقديمها بشكل مناسب للمستفيد، ويمتد دوره ليشمل مراقبة المعلومات العلمية والتقنية المنتجة من قبل مؤسسات المعلومات

ودراستها وتحليلها وتقييم كفاءتها وإتاحة استخدامها عبر الأدوات والوسائل المعروفة في مجال العمل المكتبي

ب- المدرب : وهو الشخص المسؤول عن تدريب المستفيدين على كيفية استخدام الأدوات والوسائل المساعدة في الوصول للمعلومات بأقصر وقت واقل جهد ممكن ، وهكذا يكون واحد من أهم أدواره مساعدة المستفيدين بغض النظر عن مستوياتهم وتخصصاتهم على امتلاك الأدوات الجديدة التي تسهل عليهم الوصول إلى المعرفة ، وقد يمارس المدرب دوره هذا من خلال الاتصال المباشر والآني بالمستفيد أو من خلال تنظيم الدورات التدريبية والتعريفية لهم وحسب البرامج التي تضعها مؤسسة المعلومات.

ت- الوسيط : وهو الشخص الذي توكل إليه مهمة قيادة العملية البحثية لصالح المستفيد ، خاصة في مراحلها الصعبة ، التي تتطلب بناء استراتيجية بحث منظمة تكفل الوصول إلى المعلومات بالدقة والسرعة المطلوبتين . ويعبر هذا المفهوم عن دور أخصائي المعلومات الحقيقي، فالوسيط في واقع الآمر هو مستفيد مفوض للعمل بجمع المعلومات واختيار المناسب منها لصالح مستفيد آخر (المستفيد) ، وهذا الدور يعتبر غاية في التعقيد ، حيث يتطلب إلى جانب الخبرة في إدارة العملية البحثية والمعرفة الواسعة بمصادر المعلومات، تفهم لحاجات المستفيدين والقدرة على الاستنتاج بالطريقة التي يفكر بها المستفيد ،

فضلا عن خبرة وإطلاع في مجالات موضوعية متعددة، في الغالب معظمها يكون خارج تخصصه الموضوعي.

ث- أخصائي المراجع : تعتبر طريقة توزيع وتنظيم المعلومات في بعض أنواع المصادر المعروفة باسم الكتب المرجعية ، والتي تستخدم بهدف الحصول على معلومات محددة ومباشرة عن شئ ما ، وعادة ما توفر تلك الكتب وسائل وأدوات مساعدة ، مثل (الكشافات ، الفهارس ... الخ) والتي قد تكون منفصلة أو متصلة بها لتمكن المستفيد من الوصول إلى المعلومات دون الحاجة إلى قراءة الكتاب المرجعي بالطريقة التسلسلية التقليدية التي تقرأ فيها الأنواع الأخرى من مصادر المعلومات . وعادة ما يجهل المستفيد قليل الإطلاع كيفية استخدام هذه الوسائل أو الطريقة التي استخدمت في تنظيم المعلومات بين ثنايا الكتاب المرجعي ، والأهم من هذا أن معظم المستفيدين يجهلون أين يمكن أن يجدوا المعلومات المطلوبة ، وأي المراجع الأفضل من بين ما متوفر منها. كل هذه الأسباب توجد حاجة ملحة لتوفير مساندة يقدمها موظف المراجع للمستفيد لتخطي تلك الصعوبات، ويفترض بهذا الشخص أن يمتلك الخبرة الكافية في استخدام المصادر المرجعية ، ومعرفة بالمحتوى الموضوعي لها وطريقة تنظيمها وترتيبها للمعلومات، ليتمكن من خدمة المستفيد في الحصول على المعلومات المناسبة بالحد الأدنى من الجهد والوقت.

ب/الوساطة في مجال البحث الآلي:

أن الوساطة في مجال البحث الآلي عن المعلومات كانت قد بدأت بشكل بارز مع بداية خدمات البحث بالاتصال المباشر، التي تعتمد على شبكات الاتصالات للوصول إلى المعلومات ، وكما هو معروف فان هذه الخدمات كانت تقدم من خلال مؤسسات تجارية مقابل رسوم اشتراك، يحتسب جزء كبير منها على أساس الوقت المستنفد في إجراء عملية الاتصال . من هنا كانت مؤسسات المعلومات التي لديها اشتراك بهذه الخدمات تحرص دائما على أن تكون الرسوم المالية المترتبة عليها في حدود ميزانياتها، بغض النظر عما إذا كانت مؤسسات المعلومات ستتحمل نفقات البحث بشكل كامل أو إنها ستمرره إلى المستفيد.

ففي الحالة الأولى سيكون لدى هذه المؤسسات حافز قوي لضبط النفقات والإصرار على أن يتولى أحد موظفيها مسؤولية إجراء البحث بالاتصال المباشر، والذي تتوسم فيه الخبرة والدراية في إدارة العملية البحثية بما يحقق هدفها في ضبط النفقات إلى أدنى حد ممكن ، أما في الحالة الثانية فقد يفضل المستفيد أن يستفيد من مهارات الوسيط وسرعته في إنجاز البحث لضغط النفقات المترتبة عليه شخصيا نتيجة الربط بالاتصال المباشر. فكل دقيقة إضافية يمضيها الباحث بالاتصال المباشر على الحاسوب المضيف تعني زيادة في رسوم الاتصال، وعلى هذا الأساس شجعت مؤسسات المعلومات الأبحاث السريعة والفاعلة

التي تنجز من قبل الوسطاء خاصة أولئك الذين يمتلكون الخبرة والمهارة المكتسبة من الممارسة الفعلية لعمليات البحث، بدلاً من الارتباك والتردد والبحث البطيء التي يجريها المستفيدون العرضين، الذين قد لا يكونوا عارفين بقواعد البيانات أو وظائف نظم استرجاع المعلومات.

وتجدر الإشارة إلى أن الوساطة في مجال البحث بالاتصال المباشر أسندت منذ البداية إلى المكتبيين، حيث يعود سبب تفضيله على غيره من المتخصصين للقيام بدور الوساطة إلى المبررات الآتية

أ- كانت المكتبات من مؤسسات المعلومات السباقة في الحصول على خدمات البحث بالاتصال المباشر، لخدمة المستفيدين منها ، ومن البديهي وكما أشير سابقا أن توكل مهمة الوساطة في إجراء العملية البحثية لموظف المكتبة .

ب- كانت معظم قواعد البيانات المتاحة بالاتصال المباشر قد استخدمت ذات المفاهيم البحثية التي شاع استخدامها في مجال العمل المكتبي التقليدي ، مثل عناصر البحث البولياني ولغات الأوامر ، وحتى مع وجود اختلاف وان كان بسيطاً فيما بين البحث بالأساليب التقليدية والبحث الآلي ، إلا أن تعلم لغات الأوامر وعناصر البحث البولياني كانت بالنسبة للمكتبين جزء من وظيفتهم.

ت- كانت معظم قواعد البيانات المتاحة تجاريا من خلال البحث بالاتصال المباشر تقدم في العموم معلومات ببليوجرافية ولهذا وقعت مباشرة ضمن هيمنة المكتبة كمؤسسة والمكتبين كأشخاص.

ونستطيع القول إن ارتباط هذا الدور بالمكتبيين قد منحهم إحساسا بالحماس لتقبل إدخال التقنيات الحديثة إلى مجال العمل المكتبي، ولقد آمنوا بان هذه التقنيات ستحسن من قدراتهم المهنية، كذلك ستساعد على رفع منزلة المكتبة والمكتبي في نظر الباحثين والمستفيدين.

إلا أن النمو المتزايد في بحث المستفيدين بأنفسهم شكل خطراً طويل الأمد لهذا الإحساس المتعالي بالقدرات المهنية ، فدخول المستفيد إلى مجال البحث المباشر والتخلي تدريجياً عن دور الوسيط ، قد يمس مواقف حساسة كثيرة تتراوح بين الشك بإمكانية المستفيد في إجادة تعقيدات البحث ، إلى الخوف من مستقبل لا تعود فيه حاجة لأخصائي المعلومات أبدا. خاصة بعد تطور وسائل النشر الإلكتروني وظهور تقنية الأقراص المدمجة وتنامي الشبكة العالمية للمعلومات (الإنترنت) مع ما يرافق ذلك من سهولة استخدام تقنية الحواسيب بالنسبة للمستفيد. كل هذا أوجد شعوراً لدى المكتبين بان هناك خطر يهدد وجود المهنة المكتبية عموماً ودور الوسيط خصوصاً ، على فرض أن الإنترنت ستكون أعظم من أي مكتبة في العالم فيما تحتويه من معلومات، وطالما كان بإمكان كل مستفيد أن يكون لديه محطة عمل في منزلة أو

مقر عمله يتصل من خلالها بالشبكة ، فلا تعود هناك حاجة لديه لزيارة المكتبة والاستفادة من خدماتها أو الاستعانة بخبرة الوسطاء في إنجاز العمليات البحثية .

ونعتقد أن الشعور بالخوف هنا قد لا يكون مبرر طالما كان اختصاصي المكتبات قادرين على التحول بسرعة إلى متخصصين متمكنين في البحث وتطوير خبراتهم ومهاراتهم للدرجة التي تفوق قدرات المستفيدين في التعامل بكفاءة مع كل التطبيقات الحاسوبية وواجهات البحث ومواقع ومحركات شبكة الإنترنت.

ويمكن الإقرار بان الحاجة إلى الوسيط في مجال البحث الآلي عن المعلومات ستبقى قائمة طالما يظل أكثر فاعلية في إنجاز العمليات البحثية، ويتمتع بمهارات خاصة تمكنه من تحقيق نتائج بحث أفضل من تلك التي يتمكن المستفيد من تحقيقها بغض النظر عن المهارات والفروق الفردية بين مستفيد و آخر.

وهناك مجموعة من العوامل ذات الصلة بدور الوسيط في إجراء البحث الآلي المفوض، ويحتمل أن يكون لها تأثير في كفاءة نتائج العمليات البحثية.

- **التخصص والاهتمام الموضوعي:**

بشكل عام يمكن النظر إلى تباعد التخصص الموضوعي للوسيط عن موضوع البحث أو تخصص المستفيد باعتباره عاملاً مؤثراً في كفاءة الاسترجاع للاعتبارات الآتية :

أ- درجة توافق البنية الشخصية (الموضوعية) للوسيط مع تلك الخاصة بالمستفيد.

ب- مدى تعقد الموضوع المتداول ومدى إمكانية التعبير عنه بدقة.

ت- معرفة الوسيط بالمحتوى الموضوعي لنظام استرجاع المعلومات الذي يعمل عليه.

ث- عندما يفشل الوسيط في فهم العلاقة الموضوعية الرابطة بين المصطلحات أو يعجز عن التمييز بين المصطلحات الرئيسية والمصطلحات الفرعية.

ج- إدراك الوسيط لأهداف المستفيد ودوافعه البحثية والموضوعية.

ح- لغة الاسترجاع التي يستخدمها المستفيد للتعبير عن موضوعه البحثي.

- **التواصل مع المستفيد:**

يجد الكثير من المستفيدين صعوبة في صياغة طلب بحث يصف بشكل دقيق حاجتهم إلى المعلومات، ويجد البعض منهم صعوبة اكبر في التعبير عن هذه الحاجة لشخص آخر، وكما هو معروف فان نجاح

البحث يعتمد بشكل أساسي على وضوح وتكامل الطلب الذي ينبغي أن يكون تقديراً تقريبياً معقولاً للحاجة إلى المعلومات، فكلما زاد التباين بين الطلب المقدم والحاجة إلى المعلومات انخفضت احتمالات نجاح البحث، ولقد أثبتت بعض الدراسات أن هناك علاقة مباشرة بين فشل البحث وطريقة التفاعل غير الملائم بين الوسيط والمستفيد.

فعندما لا يكون الوسيط متخصصاً في ذات المجال الموضوعي للمستفيد أو قريباً منه، يكون بحاجة ماسة إلى إجراء المقابلة وبشكل مباشر مع المستفيد والتي يكون في مقدمة أهدافها التعرف بشكل دقيق على موضوع البحث وتداخلاته، وهذه المقابلة هي تعبير دقيق عن معنى التواصل مع المستفيد. ولكي يتمكن الوسيط من تحقيق نتائج بحث تتناسب مع الاحتياجات الموضوعية للمستفيد، لابد من أن يتعرف أولا على تلك الاحتياجات ، لكن المشكلة في هذا الجانب تكمن في أن المستفيد قد يفشل في التعبير عن حاجاته الموضوعية كما هي لشخص آخر، ومن المهم في مجال البحث عن المعلومات التمييز بين الاحتياجات إلى المعلومات والتعرف على تلك الاحتياجات أو التعبير عنها، فنظم استرجاع المعلومات لا يمكن لها أن تستجيب إلى احتياجات المستفيدين إلى المعلومات كاحتياجات، وإنما تستجيب فقط لإعرابهم عنها بمصطلحات ومفاهيم بحثية محددة ، لكن ليس من

السهل دائما على صاحب الحاجة إلى المعلومات أن يعرب عنها بوضوح ودقة إلى الشخص الذي يتولى مهمة البحث عن المعلومات.

خلاصة القول أن الاتصال بين الوسيط والمستفيد يمكن أن يحقق نتائج جيدة في توجيه العملية البحثية الاتجاه الذي يحقق رضا المستفيد، وهنا يحتاج الوسيط إلى جانب خبرته في مجال البحث إلى مواهب أخرى لها علاقة بالدراسات النفسية والتي يمكن أن تتطور مع الجهد المثابر للوسيط في تسجيل الخبرات السابقة ومتابعة سلوك المستفيدين قبل العملية البحثية وأثناءها وبعد إنجازها.

7/ دور المستفيد في كفاءة استرجاع المعلومات:

المستفيد بوصفه صاحب الحاجة الفعلية للمعلومات يؤدي دورا مهما في توجيه العملية البحثية، ومن المؤكد إن له تأثير مباشر على كفاءة الاسترجاع الآلي للمعلومات. فالعملية البحثية في مراحلها المتقدمة تبدأ أولاً بشعور المستفيد بأنه بحاجة إلى المعلومات لتحقيق أهدافاً معينة لديه ، وتنتهي بشعور المستفيد ذاته بالرضا عن النتائج التي توصل إليها، وبين البداية والنهاية هناك جوانب متعددة ترتبط بسلوك وقدرات ومهارات المستفيد في مختلف مراحل البحث، والتي يتوقع لها أن تكون ذات تأثير مباشر على الكفاءة.

من ناحية أخرى فان فاعلية وكفاءة أي نظام لاسترجاع المعلومات يتم قياسها بالاعتماد على أحكام المستفيدين وإذا استثنينا اعتبارات الجودة من عوامل التقييم، فأن قدرة النظام على إشباع حاجات المستفيدين من المعلومات بأقل جهد وبأقصر وقت، تعتبر مؤشرات موضوعية على كفاءة النظام.

وتجدر الإشارة هنا إلى أن الحكم على كفاءة النظام قد تكون مسالة أكثر تعقيدا من الحكم على كفاءة الاسترجاع، فالمستفيد معنّي بالدرجة الأساس بنتائج العملية البحثية الممثلة بعدد التسجيلات المسترجعة وملاءمتها لمطلبه البحثي، والتي يمكن قياسها باستخدام مقاييس الاستدعاء والدقة، ويعتقد الباحث أن المستفيد يجب أن لا يكون الحلقة الوحيدة في عملية التقييم عندما يتعدى الأمر موضوع قياس كفاءة الاسترجاع إلى موضوع قياس كفاءة النظام ككل، والسبب في ذلك يرجع إلى:

أ- أن أحكام المستفيدين تصلح للتعرف على مستويات الدقة من خلال الكشف عن حالة التوافق ما بين التسجيلات المسترجعة ومتطلباتهم البحثية.

ب- ما هو متوافق لمستفيد ما قد لا يكون كذلك لمستفيد آخر في الموضوع ذاته

ت- ما يمكن أن يعتبره المستفيد متوافقاً في وقت ما قد لا يجده كذلك في وقت آخر.
ث- مفهوم الدقة من وجهة نظر المستفيد في الغالب يكون نسبي تؤثر فيه العوامل الخارجية المسيطرة على قراراته.
ج- تتأثر أحكام المستفيدين على دقة نتائج العملية البحثية ذاتها باختلاف اتجاهاتهم الموضوعية.

وعموماً يمكن النظر إلى دور المستفيد في توجيه العملية البحثية من خلال دراسة سلوكه قبل العملية البحثية و أثناءها وبعد إنجازها للتعرف بدقة على كل نشاط أو فعل يصدر عنه و يحتمل أن يكون له تأثير على كفاءة الاسترجاع.

وبشكل عام يمكن القول إن فشل المستفيد في تنظير حاجته إلى المعلومات أو قصوره في التعبير عنها سيشكل العقبة الأولى التي تقف بطريق المستفيد للحصول على استجابة دقيقة واسترجاع تسجيلات (وثائق) مناسبة لإشباع حاجاته إلى المعلومات.

و فيما يأتي مجموعة من العوامل التي يتوقع لها إن تكون ذات تأثير في كفاءة عملية البحث ولها صلة وثيقة بالمرحلة التي تتم فيها صياغة طلب البحث، وعلى فرض إن المستفيد يقوم بالعملية البحثية بدون وساطة:

أ- درجة تعقيد الطلب: إذا كان الطلب بسيطاً وينطوي على عدد قليل من الأوجه، أمكن صياغة طلب البحث بالطريقة التي تؤدي إلى حصول المستفيد على نتائج بحث جيدة. وكلما زادت العلاقات بين المصطلحات البحثية، تصبح عملية صياغة الطلب أكثر تعقيدا.

ب- مرونة لغة التكشيف : ينبغي أن يوفر النظام لغة تكشيف لها مرونة في التعبير عن موضوع الاستفسار بمستوى مناسب من التخصص. وكذلك ينبغي أن تكون اللغة مشتملة على العلاقات التفريعية والارتباطات الموضوعية بين المفاهيم والمصطلحات البحثية في محتوى قاعدة البيانات المبحوثة.

ت- التغطية المصدرية لقاعدة البيانات : عندما تكون التغطية المصدرية لقاعدة البيانات في مجال الموضوع الواحد شاملة، تصبح لغة التكشيف غير دقيقة أو ضعيفة في الكشف عن العلاقات الزائفة والغامضة بين المصطلحات، بسبب التباين النوعي بين المصطلحات المستخدمة في الأنواع المختلفة لمصادر المعلومات.

ث- الخبرة العملية والإمكانيات اللغوية للمستفيد: كلما تقاربت لغة المستفيد مع لغة محتوى قاعدة البيانات المبحوثة، كلما أمكن صياغة الطلب بسهولة وسرعة كبيرتين، مع الأخذ بنظر الاعتبار الخبرة العملية السابقة في تنفيذ العمليات البحثية.

ج- دعم نظام استرجاع المعلومات : على تقديم وظائف مساندة للمستفيد في التحكم بالمصطلحات من خلال البتر واستكمال الحروف المفقودة وتضمين اللواحق اللغوية والتدقيق الإملائي للمصطلحات الموضوعية.

كما أن هناك عوامل تتعلق بتنقية المستفيد لنتائج البحث، حيث يستجيب النظام لطلب المستفيد باسترجاع مجموعة من التسجيلات التي تمثل جزء من محتوى قاعدة البيانات التي تم اختيارها من قبل المستفيد، ويفترض بهذه التسجيلات أن تلبي حاجات المستفيد إلى المعلومات، والتي عبر عنها باستخدام المفاهيم والمصطلحات الموضوعية عند صياغته طلب البحث. وسبقت الإشارة إلى أن هناك أربع مستويات يمكن على أساسها تقسيم نتائج العملية البحثية، حيث يمكن أن يتم استرجاع مجموعة من التسجيلات تتوافق تماماً مع موضوع البحث، و أخرى لها علاقة قريبة بالموضوع، وثالثة لها علاقة بعيدة بالموضوع ، والاحتمال الأخير هو أن يتم استرجاع مجموعة من التسجيلات ليست لها علاقة بالموضوع

ومما لاشك فيه أن نجاح عملية التنقية تتأثر مباشرة بمستوى بيان الطلب، الذي يتم على أساسه التنبؤ بالاتصال بالموضوع . فإذا كان بيان الطلب تعبيرا قاصرا عن حاجة المستفيد من المعلومات، فإنه من الممكن أن تستبعد تسجيلات يمكن أن يراها المستفيد ذات صلة

بالموضوع . فالدقة إذن في بيان الطلب وفي قدرة المستفيد على تفسير احتياجاته إلى المعلومات هي من أهم العوامل المؤثرة في نجاح عملية التنقية أو فشلها. وتزداد أهمية هذه العملية لارتباطها المباشر بمقاييس الاستدعاء والدقة المستخدمة في تقييم نظم استرجاع المعلومات، اللذين يتم احتسابها على أساس عدد التسجيلات المسترجعة لمطلب بحثي معين، وعدد التسجيلات التي يجدها المستفيد ذات صلة بموضوع البحث، و هناك مجموعة من المشكلات التي يتوقع أن يكون لها تأثير على كفاءة الاسترجاع وذات صلة مباشرة بدور المستفيد في تنقية المخرجات ما يأتي:

أ- أحكام المستفيد على نتائج عملية الاسترجاع تخضع لدوافع شخصية غير معلنة.

ب- في قواعد البيانات الببليوغرافية التي تقدم مستخلصات البحوث فقط قد لا يتمكن المستفيد، خاصة قليل الخبرة من اكتشاف التسجيلات الملاءمة لقلة المعلومات المتوفرة في المستخلص

ت- عندما تختلف لغة قاعدة البيانات عن لغة المستفيد تصبح عائقاً أمامه للحكم بدقة على مدى صلاحية التسجيلات المسترجعة

ث- ضيق وقت المستفيد في الغالب لا يمكنه من استطلاع التسجيلات المسترجعة كافة خاصة عندما يكون عددها كبير إذ يعمل على الإسراع في عملية الاستطلاع أو تجاهل استطلاع عدد من التسجيلات

ج- في بعض الأحيان يكون لدى المستفيد قرار مسبق بعدد التسجيلات التي يرغب في الحصول عليها.

ح- الكلفة المادية لعمليات البحث قد تحدد المستفيد أحيانا بقبول عدد محدود من التسجيلات المسترجعة والاستغناء عن مجموعة أخرى حتى مع إقراره بملاءمتها.

8/ اختيار قاعدة البيانات:

بعد أن ينجز المستفيد صياغة طلب البحث باختيار المفاهيم والمصطلحات الموضوعية التي يراها مناسبة للتعبير بدقة عن موضوعه، وتحديد نوع العلاقة التي تربط فيما بين هذه المفاهيم، قد يترك العمل اللاحق للوسيط أو يقوم به بنفسه، وغالبا ما تكون المرحلة اللاحقة لصياغة طلب البحث اختيار قاعدة البيانات، التي يفترض بها أن تحقق له أفضل النتائج البحثية، وكان هذا القرار منذ ثلاثين عاماً بسيطاً نسبياً عندما كان عدد قواعد البيانات المتاحة من خلال نظم

استرجاع المعلومات المعروفة قليل للدرجة التي يتمكن المستفيد أحياناً من البحث في معظم هذه القواعد مرة واحدة ومع الزيادة الكبيرة في عدد قواعد البيانات المتاحة بالاتصال المباشر أو المنتجة على الأقراص المدمجة وتعدد نظم استرجاع المعلومات ، أصبحت مسألة اختيار قاعدة البيانات تأخذ مساحة اكبر من اهتمام المستفيد ، وتشكل عقبة في طريق تحقيق نتائج جيدة للعمليات البحثية.

ولكي ندرك حجم هذه المشكلة نشير إلى أن عدد قواعد البيانات الببليوجرافية كان عام 1987 (3288) قاعدة متاحة تجاريا من خلال البحث بالاتصال المباشر عن طريق (597) مركز خدمة معلومات في الولايات المتحدة وأوربا الغربية، منها على سبيل المثال مركز خدمة دايلوك الذي تأسس عام 1960 بقاعدة بيانات واحدة ليصل عدد القواعد فيه عام 1988 إلى أكثر من (155) قاعدة بيانات تغطي مختلف المجالات الموضوعية، ثم ازداد عددها ليصل إلى من (600) قاعدة بيانات في نهاية عام 1999، وفي المقابل كان هناك نمو أكبر في قواعد البيانات على الأقراص المدمجة حيث كان عددها عام 1991 يصل إلى (2200) قاعدة موزعة على مختلف الموضوعات العلمية والإنسانية والعامة ، وفي عام 1999 حدث تطور مهم وزيادة كبيرة في إنتاج قواعد البيانات على الأقراص المدمجة ليزداد عددها إلى من (16000) قاعدة موزعة بنسب متباينة على مختلف المجالات على

الأقراص المدمجة، في التخصصات العلمية خاصة، والتي يتم استجوابها باستخدام نظام استرجاع المعلومات.

من جانب أخر فان طريقة تناول الموضوع ونوع المخرجات التي تقدمها قاعدة البيانات تشكل عامل اختلاف آخر يضاف إلى ما تقدم ، وعليه فان موضوع اختيار قاعدة البيانات قبل إجراء العملية البحثية سيكون له تأثير على النتائج النهائية للبحث، ومن الأسباب التي تدعو إلى اختيار قاعدة البيانات هي:

أ- **التغطية المصدرية** : في بعض قواعد البيانات تكون التغطية شاملة لكافة أنواع مصادر المعلومات من كتب ودوريات ووقائع مؤتمرات ورسائل ...الخ ، في موضوع محدد ، والبعض الآخر يتخصص في نوع محدد من المصادر ، على سبيل المثال قاعدة بيانات بالرسائل الجامعية أو بالدوريات أو أي نوع آخر، ولكل من هذه الأنواع طريقته في تناول الموضوع التي قد تتوافق أو لا تتوافق مع توجهات المستفيد.

ب- **التغطية الزمنية** : في أحيان كثيرة تكون التغطية الزمنية لمحتويات القاعدة محدد بمدى زمني يطول أو يقصر حسب كمية النتاج الفكري المنشور في مجالها الموضوعي.

ت- التغطية اللغوية : قد تشتمل القاعدة على مصادر معلومات منشورة بلغات مختلفة أو قد تتخصص في لغة معينة.

ث- التخصص والشمول : بعض قواعد البيانات تحتوي على معلومات متفاوتة بدرجة تخصصها، فقد تتخصص قاعدة بيانات طبية مثلاً بفئة عمريه أو بحدود نوعية معينة، أو قد تكون شاملة

ج- المعالجة الموضوعية :الموضوع الواحد قد يتم تناوله من وجهات نظر متعددة ، وقد تكون قاعدة البيانات شاملة في تغطيتها للموضوع لوجهات النظر كافة، أو قد تقتصر على نوع محدد من المعالجة.

ح- نوع المخرجات : تختلف قواعد البيانات في الموضوع الواحد بنوع المخرجات التي تقدمها للمستفيد في نهاية العملية البحثية، ما بين معلومات ببليوجرافية، أو ملخصات أو النصوص كاملة.

وتبقى المشكلة في احتمال جهل المستفيد للتفاصيل الفنية والموضوعية لقواعد البيانات أو حتى مسميات هذه القواعد خاصة بالنسبة لقليلي الخبرة منهم.

الفصل الثالث

قواعد البيانات: نموذج لنظم استرجاع المعلومات الآلية

مقدمة:

تمثل المعلومات أهمية كبيرة للفرد والمجتمع، وتتراوح حاجة الفرد للمعلومات من حين لآخر فقد يحتاج إلى معلومات بسيطة ويستطيع العثور عليها بالرجوع إلى مرجع محدد، وقد يتجاوز الأمر ذلك فى أحيان أخرى بحيث يتطلب الرجوع إلى مجموعة من المصادر اللازمة للرد على استفسار المستفيد وتلبية احتياجاته من المعلومات. وفى ظل الفيضان الهائل من المعلومات الذى أصبح يتميز به عصر المعلومات نتيجة الزيادة المطردة فى كم المنشور وتعدد جهات ولغات نشره، هذا إضافة إلى تعقد التخصصات الموضوعية وتداخلها كل ذلك أدى إلى ضرورة وجود وسيلة تساعد المستفيد على الوصول إلى مطلبه من المعلومات وسط هذا الخضم الهائل من المعلومات.

هذا وتعتبر قواعد البيانات من أحدث الأساليب المعاصرة لتخزين واسترجاع المعلومات في تطبيقات التجهيز الآلي للبيانات. وبصفة خاصة في مجال المكتبات والمعلومات حيث يتوقع زيادة أهميتها في السنوات المقبلة وتحويلها إلى أنظمة إدارة المحتوى الرقمي لمواجهة الكميات الهائلة من المعلومات.

1/ قواعد البيانات:

تطورت قواعد البيانات بظهور وتطور الحاسبات الالكترونية، للإمكانات التي تتمتع بها الحاسبات في اختزان واسترجاع البيانات، كما هو واضح فإن مصطلح قاعدة بيانات(Data Base) يشير للحفظ والسيطرة على كم مهول من البيانات لغرض استرجاعها عند الطلب وذلك حسب رغبة المستفيد.

قبل ظهور قواعد البيانات كانت البيانات والمعلومات – بأشكالها المختلفة – تخزن في المكتبات على الرفوف أو داخل الخزانات أو أدراج فى صناديق الفهارس، وتأخذ حيزا مكانيا كبيرا وتصعب السيطرة عليها والاسترجاع من خلالها في بعض الأحيان، كما أنها عرضه لعوامل التلف والسرقة وغيرها. ثم أصبحت الحاسبات الالكترونية مستودعات تخزينية منظمة ويسهل معها اختزان واسترجاع كميات مهولة من البيانات وفى نفس الوقت. وهذه المستودعات بها ذاكرة هي ذاكرة الحاسب الالكتروني.

1/1 البيانات(Data):

يمكن ان تكون البيانات حروف أو أرقام أو رمز أو إشارات أو كلمات أو خليط منها جميعا أو بعض منها. أو يمكن أن تكون بعض

الجمل الناقصة غير المنظمة ليس لها دلالات أو معنى محدد ولا يمكن اعتمادها فقط لفهم شئ أو اعتماده إلا بعد استكمال المعنى. كما يمكن وصفها بأنها عبارة عن حقائق و أفكار وآراء تصف حدثا معينا ولكن دون إجراء أي تعديل أو تفسير أو مقارنة، ولذلك تعتبر البيانات هي المواد الخام التي تشتق منها المعلومات.

2/1 المعلومات:

عبارة عن مجموعة من البيانات يتم ترتيبها وتنظيمها ومعالجتها بصورة تعكس مفاهيم ومعان تؤدى إلى الإجابة عن الاستفسارات وإلى توضيح المعاني والجمل، وبالتالي تؤدى إلى التعلم والمعرفة وكسب المهارات، فهي الحقائق عن الأشياء والأماكن والطبيعة والأشخاص وكل شئ في الحياة.

كما يمكن تعريفها بأنها عبارة عن بيانات ولكن تمت معالجتها ليستفاد منها في عملية اتخاذ القرار، أي إنها البيانات التي خضعت للتحليل والتفسير والمقارنة.

3/1 تعريف قاعدة البيانات:

لقد أصبح مصطلح قاعدة البيانات شائع الاستخدام في أوائل السبعينيات من القرن العشرين، وفى السنوات التالية لذلك التاريخ

أصبحت قواعد البيانات واسعة الانتشار وتزايدت اهميتها يوما بعد يوم، حتى أصبح بنائها وتطويرها أحد الأنشطة الأكثر أهمية في مجال تصميم نظم المعلومات المرتبطة بالحاسب الالكتروني.

ولقد ورد في الإنتاج الفكري عدة تعريفات لقاعدة البيانات ومهما تنوعت هذه التعريفات إلا أنها جميعا تتفق على أن قاعدة البيانات هي عبارة عن ملف file مكون من مجموعة من التسجيلات Records المتصلة فيما بينها، هذه التسجيلات تضم مجموعة من الحقول fields وكل حقل من هذه الحقول يتضمن البيانات.

أو هي عبارة عن" مجموعة المعلومات أو البيانات المتصلة، ذات العلاقة المتبادلة فيما بينها المخزنة بطريقة نموذجية ودون تكرار".

وفى المعجم الموسوعي لمصطلحات المكتبات والمعلومات، فإن قاعدة البيانات يى" مجموعة منظمة من التسجيلات المقروءة آليا، وتحتوى على واحد أو أكثر مما يلى: توصيفات ببليوجرافية، أو معلومات في موضوعات محددة، أو بيانات(في شكل ارقام، أو معلومات في نص). وهذه البيانات تكون في العادة ذات بنيان ويتم إنتاجها طبقا لمعايير متسقة".

4/1 مكونات قاعدة البيانات:

تتكون قاعدة البيانات من الآتى:

1- البيانات: وهى من أهم مكونات نظم قواعد البيانات، وهى العنصر المركزي الذى تحيط به العناصر الاخرى.

2- **الحقول:** هي مواقع محددة، ومعرفة ذات دلالة ومعنى تستخدم لإدخال البيانات فيها، لتكون فى النهاية معلومات ذات معنى ودلالة عند الطلب والاسترجاع. والحقوق يمكن اعتبارها أماكن خزن وحفظ البيانات المدخلة، ولكن وفق آلية ونظام علمي ومنطقي معين ومعروف للحاسب الإلكتروني وبرمجياته المتعددة للسيطرة على البيانات وإنشاء ملفات متعددة عند الحاجة والطلب.

والحقول أنواع متعددة:

- **من حيث الحجم نوعان من الحقول:**

أ- الحقول الثابتة (Fixed Fields) والتي عادة تستخدم للبيانات ذات الأطوال المحددة مسبقا مثل رموز البلدان واللغات ...إلخ.

ب-الحقول المتغيرة الطول(Variable Length Fields) وهذه ملائمة لمختلف أشكال البيانات مثل العنوان، المؤلف، المستخلص.

- **من حيث التفريع يوجد نوعان:**

أ- الحقول الأولية (Elementary Field) وهى الحقول التي تحتاج بياناتها إلى التفريع مثل حقل العنوان، الوصف المادي.

ب-الحقول المتفرعة (Sub Fields) وهى الحقول التي تحتاج بياناتها إلى التفريع، مثل بيانات التأليف، النشر. وهذه أصغر ما في الحقول حيث يتم تجزئة بيانات الحقل الواحد إلى مجاميع أو فئات، وتستخدم الرموز (Codes) للفصل بين الاجزاء أو الفروع، وتختلف في شكلها بين نظام أو برنامج وآخر.

ومهما كان نوع وطول الحقل فيجب أن يكون له اسم مميز ينفرد به، وهذا الاسم قد يكون أرقاما أو حروفا أو الاثنين معا. وللحقول مواصفات تحددها طبيعة البيانات التي ستدخل في الحقل.

3- التسجيلات (Records): ومفردها تسجيلة وتمثل وحدة واحدة، أو شكل محدد، أو مصدر معلومات محدد. وتحتوى التسجيلة على جميع المعلومات الخاصة بالوعاء، سواء كانت كتاب أو مقال في مجلة أو رسالة جامعية. إذا التسجيلة عبارة عن مجموعة من الحقول. والتسجيلات أنواع وأحجام مختلفة، إذ يعتمد نوع وحجم التسجيلة على أنواع الحقول المستخدمة فيها، وعلى أطوالها وعددها، والتي يعود

اختيارها إلى نوع البرنامج المستخدم في قاعدة البيانات، وإلى نوع وطبيعة البيانات التي ستدخل في هذه الحقول.

وقد تفكر كل مكتبة تستخدم الحاسوب في عملياتها الفنية في إعداد تركيبة خاصة بها، وهذا ممكن من الناحية النظرية، غير أن عالم الواقع أصبح فيه التقنين والتوحيد والموائمة ضرورة قصوى من أجل التعاون ومشاركة الموارد، وبالتالي بناء شبكات معلومات على كل المستويات، ولهذا برزت جهود حثيثة في عدد من المؤسسات الرائدة في خدمات المعلومات، من أجل وضع تركيبات تصلح للتطبيق في نظم معلومات متنوعة. ومن أمثلة التركيبات العالمية ما يلى:

أ/ تركيبة مارك (Machine Readable Cataloging: MARC):

بدأ العمل بها عام 1965، من خلال مشروع الفهرسة المقروءة أليا، ثم ظهور الإصدارات المختلفة (مارك1، مارك2، ثم مارك لكل أنواع أوعية المعلومات)، ومع بداية السبعينيات من القرن العشرين وبظهور الاتجاه نحو الفهرسة التعاونية، أصبح شكل مارك وسيلة لتبادل

التسجيلات الببليوجرافية، وتزامن هذا التطور في الفهرسة مع تطور الجيل الثالث للحواسيب في خاصية المشاركة في الوقت.

وهذه التركيبة مطابقة لقواعد الفهرسة الأنجلو–أمريكية الطبعة الثانية، ولذا فإن التسجيلة الخاصة لأى مادة مكتبية تلتزم بحقول الوصف الببليوجرافى المنصوص عليها في هذه القواعد ولكنها تحتوى على بيانات أخرى مثل: رقم تصنيف ديوي العشري ورقم تصنيف مكتبة الكونجرس ورؤوس الموضوعات وأرقام الوثائق ولغتها...

ويتكون نموذج تسجيلة مارك من ثلاث مكونات هى:

1/ القائد(Leader): وهو عبارة عن افتتاحية تصف التسجيلة وتسمح بالتعرف عليها ويقدم معلومات عن طولها ونوعها ومستواها الببليوجرافى ونقصد بذلك ما يلي:

- الطول(طول التسجيلة): يتكون القائد دائما من 24 حرفا أى أنه طول ثابت.
- نوع التسجيلة:كتاب، دورية، مخطوط....
- المستوى الببليوجرافى: والمقصود منه هل العمل عبارة عن كتاب كامل؟ أو جزء من سلسلة، أو دورية، أو مدخل تحليلى لجزء من عمل شامل؟

2/ الدليل(Directory): وهو بمثابة صفحة محتويات أو كشاف دقيق لمكان تواجد البيانات الببليوجرافية داخل التسجيلة، ويتكون كل مدخل لدليل تسجيلة ما من مؤشر المحتوى يسمى التاج(Tag) تحدد وتعرف الحقل أو الحقول المتغير منها، ويعمل الدليل على تسهيل استرجاع حقول مختارة أو معينة فى تسجيلة ما من تسجيلات مارك.

3/الحقول الثابتة(Fixed Fields): ويقصد بها الحقول التى تحتوى على البيانات الببليوجرافية الأساسية لوصف أوعية المعلومات شكلا ومحتوى، وتكون البيانات الببليوجرافية مصحوبة بعلامات أو إشارات تشير إلى محتويات الحقول الرئيسة والفرعية مثال ذلك:

حقل بيانات النشر يتكون من حقول فرعية هي: مكان النشر، الناشر، تاريخ النشر. ويختلف طول هذه الحقول باختلاف نوعية المعلومات التي تشتمل عليها.

ب/تركيبة التراسل المشتركة (Common Communication Format:CCF):

وهى التركيبة التى طورتها اليونسكو حيث أصدرت طبعتها الثانية فى عام 1988، ويتطابق هيكلها مع المواصفة الدولية (ISO 2709 ،

والمواصفة القياسية العربية 668). وصممت لتوفير تركيبة قياسية لثلاثة أغراض رئيسة هى:

- إتاحة تبادل التسجيلات الببليوجرافية بين مجموعات المكتبات وخدمات الاستخلاص والتكشيف.
- السماح لأي مرفق ببليوجرافى بمعالجة التسجيلات الببليوجرافية المستلمة من كل من المكتبات وخدمات الاستخلاص والتكشيف بواسطة مجموعة واحدة من برامج الحاسوب.
- تصلح أن تكون الأساس لتركيبة من أجل قاعدة البيانات الببليوجرافية لدى أى مرفق ببليوجرافى.

هذا ولكى تكون قاعدة البيانات ناجحة ومفيدة وفعالة وتؤدى الغرض الذي أنشئت من أجله يجب أن تتضمن الآتي:

1- أن تبنى القاعدة بطريقة يمكن للمعلومات أن تسترجع من خلالها أو بواسطتها بصورة فعالة ومفيدة واقتصادية.

2- أن يتم بناء وتصميم القاعدة على نوعية وكمية ودقة وصحة واكتمال وحداثة المعلومات التي تدخل في التسجيلات.

3- عدد الحقول المستخدمة وملاءمتها لطبيعة وشكل ومصادر المعلومات التي تمثلها من كتب أو مخطوطات أو دوريات أو رسائل جامعية....

4- تصميم الحقول والتسجيلات والعلاقة والربط ما بين الحقول، والتى تسهل الوصول إلى البيانات واسترجاعها.

5- يجب أن تصمم وتبنى القاعدة بطريقة تسهل على المستفيدين والمستخدمين من استرجاع معلومات مفيدة بدون جهد كبير أو إتباع طرق معقدة ومضيعة للوقت.

6- القدرات الاسترجاعية: إن نجاح أو فشل أى قاعدة بيانات يعتمد على القدرات والمواصفات التى توفرها فى تنفيذ استراتيجيات البحث السهلة والعميقة والمتشعبة فى نفس الوقت. مثلا يجب أن يكون الاسترجاع على أساس الموضوعات والربط ما بين هذه الموضوعات هو الأساس الأول والاهم ولا يكون الاعتماد على اسم المؤلف أو تاريخ النشر أولا.

7- تعدد نقاط الإتاحة: إن أهم نقطة لنجاح أي قاعدة بيانات هو إتاحة المجال أمام المستخدم للوصول إلى ما يحتاجه من معلومات عبر مداخل وطرق متعددة. فعن طريق قاعدة البيانات يمكن الوصول

إلى مصدر المعلومات من خلال أي حقل من حقول القاعدة، مثل رقم التصنيف، مكان النشر، الناشر، الرقم المعياري الدولي.

5/1 أهمية وفوائد قواعد البيانات:

لقواعد البيانات أهمية وفوائد عدة منها:

1- تقديم خدمات معلومات أفضل لأكبر عدد ممكن من المستفيدين.

2- مواجهة الزيادة المهولة فى المعلومات، ومصادرها المختلفة، وارتفاع أسعارها، يقابله تراجع فى الموارد المالية المتاحة للمكتبات.

3- توفير الوقت والجهد فى الإجراءات وتقديم الخدمات المعلوماتية.

4- توفير أرضية مشتركة للتعاون مع المكتبات ومراكز المعلومات الأخرى.

5- إتاحة قواعد البيانات على الخط المباشر.

6- توفير إمكانات متنوعة للبحث من خلال مداخل مختلفة ومنافذ استرجاع متعددة متوفرة فى قواعد البيانات.

7- السرعة فى الوصول إلى المصادر المعلوماتية.

8- الدقة فى البيانات المودعة داخل القاعدة.

9- التخلص من الورق والاختصار الشديد فى المكان.

10- الحداثة فى البيانات.

11- توفير المعلومات عند الطلب.

12- توحيد الملفات الببليوجرافية وبالتالى البيانات الموجودة فى المكتبة كلها فى موقع واحد، ويمكن للجميع الاستفادة منها بدلا من بعثرتها فى أماكن ومواقع عدة.

13- التقليل بشكل كبير من عمليات التكرار فى المعلومات والبيانات الموجودة فى أقسام المكتبة الفنية.

14- إمكانية التقاسم أو المشاركة فى الاستفادة من البيانات، والتقاسم هنا أو المشاركة فى البيانات نعنى بها إمكانية استخدام نفس البيانات المخزنة لإنتاج تطبيقات جديدة، أو الحصول على نتائج مختلفة من نفس البيانات المخزنة.

15- إمكانية تحقيق المعيارية فى التعامل مع البيانات، فقد أصبح بالإمكان الآن التعامل مع البيانات فى قواعد البيانات بسهولة ويسر، إضافة إلى إمكانية تناقل البيانات عن بعد، من تصدير واستيراد وتبادل دولى للملفات المقروءة آليا.

16- قابلية التطوير عند وجود متطلبات جديدة أو بناء نظام جديد.

17- سهولة استرجاع البيانات بالشكل الذى تفرضه رؤوية المستخدم ومتطلبات النظام.

18- مركزية التحكم والمراقبة: وذلك عن طريق تحديد جهة وحيدة مسئولة عن جميع العمليات الخاصة بإدارة قاعدة البيانات، خاصة تلك المتعلقة بصيانة وحماية البيانات وتحديد الوصول.

6/1 أنواع قواعد البيانات (Types of Data Bases):

لقد تنوعت قواعد البيانات واختلفت استخداماتها وأشكالها وطبيعة البيانات المخزنة فيها، ونماذج لغات البرمجة التي كتبت بها. حيث يمكن تخزين قواعد البيانات على أقراص مدمجة CD-Rom، ويمكن استخدامها محليا أو عن بعد. والكثير من هذه القواعد تتيح استخدام معلوماتها للمستفيدين مجانا، مثل قواعد بيانات المستخلصات، والكشافات، والأدلة وغيرها. والبعض الآخر من قواعد البيانات لا تسمح بالبحث فيها والحصول على معلومات إلا بموجب اشتراك أو مقابل مادى. هذا من حيث الاستخدام أما من حيث النوع فقد تكون قاعدة بيانات مشتركة بين أقسام المؤسسة الواحدة او بين مجموعة من المؤسسات، أو تكون قواعد بيانات خاصة بكل قسم أو بكل مؤسسة. ونظرا لكثرة هذه الأنواع وأهمية التعرف عليها، نستعرض بعض أنواع قواعد البيانات:

يمكن تقسيم قواعد البيانات إلى الأنواع التالية:

أولا: حسب طبيعة البيانات التى تحتويها:

وتقسم إلى الأنواع التالية:

1/ قواعد البيانات الببليوجرافية Bibliographical Data Bases:

وهى أقدم أنواع قواعد البيانات التي صممت بواسطة نظم الحاسبات الإلكترونية في بيئة المكتبات. وهى في واقع الحال أشبه بالكشافات التي تعمل على تحديد مواقع للبحوث المنشورة في الدوريات مع إعطاء المعلومات الببليوجرافية المتكاملة، المؤلف، العنوان، بيانات النشر، المستخلص.

ويقوم هذا النوع من قواعد البيانات بدورا مهما في مساعدة المكتبات والباحثين في اختصار الوقت والجهد والمسافات الجغرافية البعيدة، وفتحت الأبواب المغلقة أمام رصد مواقع وأماكن تواجد مصادر المعلومات المطلوبة، فهي لا تقود الباحث إلى المعلومات بشكلها النصي مباشرة، بل تعرفه بما هو منشور ومتوافر من مصادر عن المجال الذي يبحث فيه ويفتش عنه، ومن نماذج مثل هذه القواعد قاعدة أريك ERIC التعليمية، وقاعدة مدلاين MEDLINE الطبية، وقاعدة

أكريكولا AGRICOLA الزراعية، التي من أهم القواعد العالمية، التي تعمل على تحليل وتكشيف واسترجاع الإنتاج الفكري للاختصاصات المذكورة.

2/ قواعد بيانات النص الكامل Full Text Data Bases:

وهى المرحلة المتطورة لقواعد البيانات الببليوجرافية حيث أضافت البعد المهم الذى يحتاجه الباحث والمستفيد ألا وهو النص الكامل لمصدر المعلومات سواء على صيغة PDF أو HTML إضافة إلى البيانات الببليوجرافية. ويعد هذا النوع من أشهر أنواع قواعد البيانات في العالم. وهو في تزايد مستمر، بعد أن وجد الباحثون والمستخدمون بأن قواعد البيانات الببليوجرافية ليست وافية، وبعد أن توسعت القدرات التخزينية للحاسبات. ومن أمثلة هذا النوع LexisNexis Academic و JSTOR.

3/ قواعد البيانات المرجعية Reference Data Bases:

وهذا اتجاه جديد في قواعد البيانات المتاحة عبر شبكة الانترنت وعلى الخط المباشر، وتتضمن المراجع الالكترونية مثل الموسوعات والأدلة والمعاجم المعروفة بالمراجع الالكترونية (E-References)

التي أصبحت تستخدم بطريقة إلكترونية مختلفة تماما عن الشكل اليدوي التقليدي. وتضم النصوص الكاملة والربط ما بين النصوص والأشكال والرسوم والمجلدات على مستوى عالمي.

ثانيا: قواعد البيانات حسب الشكل:

ونقصد بها أنواع قواعد البيانات حسب الأوعية المسجلة عليها والتي تشمل ما يلى:

1- قواعد البيانات المخزنة في ذاكرة الحاسب الإلكتروني على الأقراص الصلبة.
2- قواعد البيانات على الأقراص المدمجة CD-Rom.
3- قواعد بيانات كاملة الحركةFull Motion DVD Data Bases .

ثالثا: قواعد البيانات حسب الإتاحة Accessibility:

1- قواعد البيانات المصممة محليا وتكون متاحة للعاملين والمستفيدين فى المكتبة ذاتها.
2- قواعد البيانات المعتمدة على البرمجيات الجاهزة ولكنها متاحة للعاملين فى المكتبة ذاتها فقط.

3- قواعد البيانات المتاحة على الخط المباشر On-Line وهى متاحة للمستفيدين ضمن الاشتراك فى خدمات هذه القاعدة بغض النظر عن المسافات الجغرافية.

4- قواعد البيانات المتاحة مباشرة عبر الانترنت، وهذه القواعد حاليا يمكن أن تشمل كافة الأنواع السابقة الذكر، والتي شملتها تقنيات الانترنت التي ألغت المسافات والحدود، وأوجدت لنا المكتبات الافتراضية Virtual Lbraries والتي تشتمل على نظم وقواعد البيانات على الحاسب الآلي الآن متاحة للجميع. ولكن أغلبها أصبحت تشترط الاشتراك للسماح بالدخول والحصول على المعلومات، والقليل منها لا يزال يقدم خدمات الإتاحة بدون مقابل Free Access Point.

رابعا: قواعد البيانات حسب التغطية الوعائية:

تغطى قواعد البيانات جميع أوعية المعلومات كالكتب والمقالات والتقارير الفنية والرسائل الجامعية وغيرها، إلا أن بعضها ينفرد بالتغطية لنوع واحد من الأوعية ، وهناك أمثلة كثيرة منها:

- قاعدة مستخلصات الرسائل الجامعية international dissertation abstract وتغطى رسائل الدكتوراه.
- قاعدة LC/Line وتغطى الكتب.
- قاعدة Nexis وتغطى الصحف بصفة رئيسية.
- قاعدة أبحاث المؤتمرات Conference Papers Index .
- قاعدة براءات الاختراع World Patents Index .

7/1 نظم إدارة قواعد البيانات Database Management System:

كلمة قاعدة تعنى تنظيم الشئ على أسس. هذا الأساس هو الذى يحدده نظم إدارة قواعد البيانات، حيث لابد من وجود نظم معينة لتنظيم وإدارة البيانات المخزنة في قاعدة البيانات وهو ما يطلق عليه اسم "نظم إدارة قواعد البيانات" وهى مجموعة من البرامج الجاهزة التي تساعد على تنظيم البيانات بطريقة تسمح بالوصول إليها بسرعة وسهولة. وتعديلها واسترجاعها بطرق شتى. وهى المسئولة أيضا عن تنفيذ جميع الوظائف المطلوبة من قاعدة البيانات. فمثلا بعد إضافة عملاء جدد لدليل الهاتف فى مصر فإنك قد تحتاج لإعادة ترتيب أسماء المشتركين أبجديا أو لترتيب عناوينهم، مثل هذا العمل يطلق علية إدارة قاعدة البيانات.

فهذه النوعية من البرامج توفر إمكانية إدخال البيانات وتخزينها وتعديلها واسترجاعها بطرق شتى. ومن ثم التحكم فى هيكل بناء قاعدة البيانات.

1/7/1 وظائف نظم إدارة قواعد البيانات:

- إضافة معلومة أو بيان جديد إلى الملف.
- حذف البيانات القديمة والتي لم تعد هناك حاجة إليها.
- تغيير بيانات موجودة تبعا لمعلومات استحدثت.
- البحث والاستعلام عن معلومة أو معلومات محددة.
- ترتيب وتنظيم البيانات داخل الملفات.
- عرض البيانات في شكل تقارير أو نماذج منظمة.
- حساب المجموع النهائي أو المجموع الفرعي أو المتوسط الحسابي لبيانات مطلوبة.

2/7/1 مزايا نظم إدارة قواعد البيانات:

أ-زيادة إنتاجية المستخدمين:

تتميز نظم إدارة قواعد البيانات بسهولة الاستخدام، وبالتالي يستطيع المستخدمون الحصول على المعلومات التي يحتاجونها بدون تكبد عناء التفاصيل الفنية.

ب- زيادة أمن المعلومات:

على الرغم من أن أقسام مختلفة قد تشارك البيانات معا، فإنه يمكن تقييد الوصول لمعلومات معينة لمجموعة مختارة من المستخدمين. حيث تساعد قواعد البيانات على تحقيق السرية الكاملة للبيانات المخزنة بها بحيث لا تتاح أية معلومات لأى شخص ليس له الحق في الاطلاع عليها.

ج- زيادة سرعة الاسترجاع:

حيث يتم تخزين جميع البيانات بكافة الأنشطة لجهة ما بطريقة متكاملة ودقيقة، وتصنيف وتنظيم هذه البيانات بحيث يمكن استرجاعها في المستقبل.

د-مراقبة قاعدة البيانات:

من أهم مميزات نظم إدارة قواعد البيانات متابعة التغيرات التى تحدث فى البيانات المخزنة، وإدخال التعديلات اللازمة عليها حتى تكون دائما فى الصورة الملائمة لاستخدامها فور طلبها.

2/ مراحل تطور قواعد البيانات الببليوجرافية:

إن المتتبع لتاريخ تطور قواعد البيانات الببليوجرافية يجد أن التقنيات المعلوماتية التي تعاقبت عليها قد أثرت عليها من عدة نواحي سواء في شكلها المادي أو إمكاناتها البحثية والاسترجاعية، أو طريقة إتاحتها، أو طريقة تحديثها، أو نوعية مخرجاتها وغير ذلك. وبالإمكان تقسيم المراحل التي مرت بها هذه القواعد اعتمادا على النواحي سابقة الذكر إلى أربع مراحل وهي:

1/2 مرحلة الكشافات والببليوجرافيات المطبوعة:

وهذه المرحلة دامت طويلا إلى أن بدأت شركات ومؤسسات المعلومات استخدام الحاسب وتقنياته فى إنتاج تلك القواعد. وهذا لا يعنى بالضرورة أن تلك الجهات قد توقفت عن إنتاج الكشافات والببليوجرافيات المطبوعة، بل إن بعضها لا زال يصدر بالشكل المطبوع والمحسب معا إلى اليوم.

وقد اعترى هذه المرحلة إذا ما قورنت بالمراحل اللاحقة بعض أوجه القصور، ومنها: صعوبة البحث النسبي وخصوصا إذا استلزم البحث استخدام عدة مجلدات ومحدودية المداخل البحثية (مؤلف، عنوان، موضوع) والإمكانيات الاسترجاعية، تأخر التحديث وصعوبة استخدامها من قبل عدة أشخاص في وقت واحد.

2/2 مرحلة القواعد الببليوجرافية المحسبة على الخط المباشر عبر خط التليفون والمودم:

وفى هذه المرحلة تم تحويل بيانات الكشافات والببليوجرافيات المطبوعة إلى شكل مقروء آليا بواسطة الحاسب (MARC)غالبا، كما تم تطوير البرمجيات اللازمة للاسترجاع. ثم أتيحت تلك القواعد على الخط المباشر عبر خطوط التليفون باستخدام أجهزة المودم.

وقد تميزت هذه المرحلة بعدة سمات نذكر منها:

1) سهولة البحث النسبية(المقصود هنا سهولتها مقارنة بمرحلة الكشافات والببليوجرافيات المطبوعة حيث تستغرق العملية وقتا أقصر بكثير من سابقتها وخصوصا إذا استلزم البحث استخدام عدة مجلدات.

2) وإمكانية البحث فى عدة سنوات مجتمعة أو خلال الفترة التى تغطيها القاعدة كاملة دفعة واحدة، يقول بيتر جاكسو (1998) Jacso " بإمكانك اليوم أن تبحث لمدة عشر دقائق وتحصل على نتيجة مشابهة لما كنت تحصل علية فى يوم أو يومين فى عهد الكشافات والمستخلصات المطبوعة"، بل الأكثر من ذلك أصبح بالإمكان البحث فى عدة قواعد مختلفة فى وقت واحد. كما هو الحال فى DIALOG مثلا.

3) استحدثت أساليب جديدة للاسترجاع مثل البحث البوليانى والبتر والتجاور والتحديد، وإضافة مداخل جديدة فشملت بذلك كامل التسجيلة الببليوجرافية مع المستخلص.

4) تنوعت المخرجات فشملت العرض على الشاشة أو الطباعة على ورق أو الحفظ على قرص للاستخدام لاحقا.وأصبح التحديث يتم من قبل المنتج مباشرة بشكل أسرع.

ورغم هذه المميزات الكثيرة برزت بعض السلبيات لهذه الفترة منها على سبيل المثال:

أ- التكلفة الباهظة، حيث شملت ثلاثة أنواع من الرسوم: رسوم الاشتراك السنوية، ورسوم الدخول على القواعد(لكل دقيقة اتصال).

وهذه تدفع للشركة المنتجة أو الموزعة للقواعد، وأخيرا رسوم الهاتف والتي تدفع لشركة الهاتف المحلية مقابل استخدام خط الهاتف للاتصال بتلك القواعد.

ب- صعوبة طريقة البحث في تلك القواعد حيث لم يكن غالبية الباحثين في ذلك الوقت قد تآلفوا مع أجهزة الحاسب وأتقنوا استخدامها بعد. ويضاف إلى ذلك أن طرق البحث في القواعد في ذلك الوقت كانت جديدة على المستخدمين. كما أنها هي في ذاتها صعبة نسبيا حيث اعتمدت على واجهة الأوامر التفاعلية والتي تدخل فيها الأوامر وتنفذ عن طريق محث الأوامر. ومن أشهر أمثلة ذلك نظام الاسترجاع في قواعد ديالوج التقليدي DIALOG CLASSIC والذى لا زال يستخدم واجهة الأوامر التفاعلية.

3/2 مرحلة القواعد الببليوجرافية على الأقراص المدمجة:

وفى هذه المرحلة تم تخزين بيانات القواعد على الأقراص المدمجة CD-Rom وإتاحتها في الأسواق كبديل عن القواعد على الخط المباشر. وقد اعتمدت القواعد في بداية هذه المرحلة على نظم استرجاع ذات واجهات تفاعلية مشابهة لتلك التي كانت في المرحلة السابقة، ويرجع السبب فى ذلك إلى أن نظام التشغيل السائد في ذلك الوقت كان

نظام DOS المبنى أساسا على واجهة الأوامر التفاعلية. بعد ذلك تطورت هذه النظم وصدر منها أنظمة استرجاع تعتمد واجهة المستخدم الرسومية التفاعلية(نظام النوافذ التى تظهر على الشاشة لتيسر للمستخدمين التعامل مع الحاسب، والتى تستخدم نظام الأيقونات والقوائم المنسدلة بدلا من إدخال الأوامر بكتابتها عن طريق لوحة المفاتيح.

وأما إيجابيات هذه المرحلة فنذكر منها:

أ- انخفاض التكلفة بشكل كبير مقارنة بالمرحلة السابقة، وأصبح بإمكان المكتبات استخدام القواعد بشكل مفتوح دون الحاجة إلى دفع أى رسوم سوى رسوم الاشتراك السنوية.وقد زاد هذا من إقبال المكتبات ومراكز المعلومات على هذا النوع من القواعد وإتاحتها للرواد بشكل أكبر، وبذلك ألف الباحثون استخدامها والتعامل معها وخصوصا مع ظهور الإصدارات ذات الواجهات الرسومية سهلة الاستخدام .

ب- إمكانية إتاحة تلك القواعد على الشبكات المحلية LAN حيث يتيح ذلك استخدامها من قبل عدة مستخدمين فى وقت واحد. كما يمكن الوصول إليها عن بعد بواسطة الشبكة المحلية أو ربطها بشبكات أكبر أو الاتصال بها عبر المودم باستخدام برامج خاصة بذلك .

ت- وفى هذه المرحلة أيضا بدأت بعض القواعد بإتاحة النصوص الكاملة لمقالات الدوريات. وقد أدت هذه التقنية الجديدة إلى خفض التكاليف بشكل كبير.

أما أهم السلبيات فقد تمثلت في:

أ- الحاجة إلى إنشاء شبكة محلية خاصة بهذه القواعد وذلك حتى يتسنى استخدامها من قبل عدة أشخاص في وقت واحد. حيث لم تكن الشبكات المحلية قد انتشرت في المكتبات بعد، فاحتاج الكثير من المكتبات إلى إنشاء شبكات صغيرة خاصة بهذه القواعد، والمشكلة هنا تكمن فى قلة المتخصصين في صيانة الشبكات في ذلك الوقت، وقد عانى الكثير من المكتبات من تعطل الشبكات في كثير من الأوقات أو توقف اتصال أحد الأجهزة المربوطة بالشبكة فجأة.

ب- أما السلبية الثانية فقد كانت تأخر التحديث، حيث يتطلب ذلك إصدار نسخة جديدة محدثة من القاعدة فى فترات منتظمة ثم إرسالها إلى المكتبات المشتركة حيث تستبدل بها النسخ القديمة.

4/2 مرحلة القواعد الببليوجرافية المباشرة عبر الانترنت:

وهذه هي المرحلة الأخيرة التي نعاصرها الآن، وهذا لا يعنى أن المراحل السابقة قد انتهت ولم تعد تقنياتها مستخدمة الآن. بل على العكس لازالت كلها موجودة ومستخدمة. ولكنها تقل تدريجيا مع مرور الأيام. هذه المرحلة تعتمد على الانترنت كوسيلة اتصال بين المنتج والمكتبات المشتركة. حيث تتيح تلك الشركات قواعدها على موقعها على الشبكة وتضع عليها بعض القيود بحيث لا يسمح بالدخول إليها إلا للمشتركين فقط.

وقد تشاركت هذه المرحلة مع المرحلة السابقة فى الكثير من المميزات، مثل سهولة الاستخدام وطريقة البحث وإمكانياته المتعددة، سرعة التحديث، إمكانية البحث فى عدة قواعد مختلفة دفعة واحدة وغيرها. وتميزت عن المراحل السابقة ببعض الأمور منها:

أ- استغلال إمكانيات وتقنيات النسيج العنكبوتي للانترنتWeb مثل النصوص الفائقة HYPERTEXT والروابط وإرسال النتائج عبر البريد الإلكتروني وغيرها.

ب- كما تميزت بكونها متاحة عبر الانترنت وذلك يتيح أمرا آخر مهم جدا وهو توسيع خدمات المكتبة خارج حدود جدرانها. حيث يستعمل

الكثير من الرواد الانترنت فى مكاتبهم ومنازلهم، وبذلك يمكنهم الدخول إلى تلك القواعد من هناك.

ت- هناك ميزة كبيرة لهذه المرحلة والتي يمكن اعتبارها مرحلة خامسة مستقلة بحد ذاتها وخصوصا مع التقنيات الجديدة للويب 2، وهى إتاحة النصوص الكاملة للبحوث والمقالات. ولك أن تتخيل رائد المكتبة وهو في بيته يبحث في القواعد ثم يقوم بتنزيل النصوص الكاملة للمقالات والوثائق التي يريد إلى حاسبه الشخصي دون الحاجة إلى الذهاب إلى المكتبة. وتجدر الإشارة هنا إلى أن بعض الشركات طبقت مؤخرا فكرة جديدة وهى ربط التسجيلات الببليوجرافية للوثائق بمواقع المجلات التي نشرت بها للحصول على النصوص الكاملة لها بدلا من إعادة طباعتها أو تخزينها على هيئة صور PDF. أضف إلى ذلك أن إتاحتها عبر الانترنت يعفى المكتبات من صيانة القواعد أو تحديثها.

وأما السلبيات فقليلة. منها الحاجة إلى ربط المكتبة بالإنترنت وهذا يحتاج إلى دفع تكاليف دورية لشركة الاتصالات أو لمزودي الخدمة. وهذه في حقيقة الأمر رغم كونها مصاريف إضافية إلا أنها قد تعد حسنة وذلك لأنها تشجع المكتبات على الاتصال بالإنترنت والاستفادة

من خدماتها. وهذا مع التقدم الحالي أصبح من الدعائم الضرورية لخدمات المكتبات كما أن من سلبياتها كذلك ارتفاع التكلفة النسبي.

الفصل الرابع

استراتيجيات البحث في قواعد البيانات

مقدمة:

لا شك أن الهدف الأساسي لبناء نظم استرجاع المعلومات يتمثل فى إتاحة هذه النظم للمستفيدين، وتلبيتها لاحتياجاتهم واستفساراتهم البحثية، وذلك من خلال استرجاع المعلومات المناسبة التي تلبى تلك الاحتياجات أو تجيب عن تلك الاستفسارات. والمعلومات فى هذا السياق تعبير عام يقصد به مخرجات تلك النظم أو نتائج عمليات البحث فيها. ونوعية تلك المخرجات عادة ما تختلف باختلاف طبيعة قاعدة البيانات نفسها وأهدافها وسمات جمهور المستفيدين منها؛ فالبيانات الببليوجرافية تعد بديلا عن الوثائق تمثل فى الأساس نتائج البحث فى قواعد البيانات الببليوجرافية، هذا في الوقت الذى تمثل فيه المعطيات والحقائق نتائج البحث فى قواعد البيانات النصية أو غير الببليوجرافية.

وعادة ما تقاس كفاءة نظم استرجاع المعلومات بقدرتها على استدعاء المعلومات المناسبة أو الصالحة أي التي تتفق وحاجة المستفيد من ناحية، وحجب المعلومات غير المناسبة أو الحد قدر الإمكان من استرجاع المعلومات غير الصالحة من ناحية أخرى، ويعبر عن الطرف الأول من تلك المعادلة بمعدل الاستدعاء، فى حين يعبر

عن الطرف الثاني بمعدل التحقيق. وبالرغم مما يكمن من علاقة عكسية بين هذين الطرفين، فإن حرص القائمين على تلك النظم والحريصين على تلبية حاجات المستفيدين يكاد ينصب على رفع معدلات طرفي المعادلة–أي التحقق والاستدعاء.

وتتوقف كفاءة نظم الاسترجاع وتحقيقها لمعدلات التحقيق والاستدعاء المنشودة على الكثير من العوامل الفكرية التي تكاد تتصل بجميع مكونات تلك النظم، وبأساسيات البحث فيها، وسبل التعامل معها، بدءا بالمستفيد ومدى قدرة على التعبير عن حاجته بدقة ووضوح، مروراً باختصاصي المعلومات وكفاءته في تحويل استفسار المستفيد إلى استراتيجية بحث تتلاءم ونظام الاسترجاع، ثم بكفاءة لغة التكشيف المستخدمة من حيث التخصيص والشمول، وانتهاء بكفاءة برمجية البحث والاسترجاع ذاتها التى تُدار بها قاعدة البيانات.

1/ برمجية البحث والاسترجاع فى قاعدة البيانات:

هي المكون الفكرى softwares الذى يعد بمثابة حلقة الوصل بين المكونات المادية hardwares وملفات البياناتdatawares ، ومن ثم فهي البرمجة المسئولة عن إدارة قاعدة البيانات، والتى تتيح للمستفيد فرصة التعامل مع جميع الملفات وبحثها، واسترجاع التسجيلات التي

تتناسب وتساؤلاته، كما تمكنه من طباعة هذه التسجيلات أو تحميلها. وتعرف هذه البرمجيات بمسميات متعددة في الإنتاج الفكري، حيث تعرف ببرمجيات البحث والاسترجاع، كما تعرف في أحيان أخرى ببرمجيات البحث فقط، وفى أحيان ثالثة تعرف ببرمجيات الاسترجاع. وتتفاوت برمجيات البحث والاسترجاع تفاوتا واضحا فى إمكاناتها وخصائصها نتيجة لتفاوت الشركات التي تنتجها، ومرد هذا التفاوت يرجع إلى حدة التنافس بين تلك الشركات وحرص كل منها على كسب السوق.

2/ استراتيجية البحث:

تعد استراتيجية البحث بمثابة التعبير الفني عن تساؤل المستفيد، أو هي ترجمة حاجة المستفيد إلى مجموعة من المصطلحات المترابطة فيما بينها بأسلوب معين يضمن الاسترجاع الأمثل للمعلومات التي تلبى هذه الحاجة من ناحية، ويحد من الاسترجاع الخاطئ للمعلومات التي تخرج عن إطار اهتمام المستفيد من ناحية أخرى. وتنطوي عملية صياغة استراتيجية البحث على مرحلتين أساسيتين هما:

أ- التحليل الموضوعي.

ب-الترجمة.

ففي مرحلة التحليل الموضوعي يقوم الباحث بالفهم الدقيق لموضوع تساؤله وما يتم البحث عنه، وترجمة هذا الفهم إلى مجموعة من المصطلحات التي تتفق ولغة التكشيف المستخدمة في قاعدة البيانات. ويرى لانكستر أن جودة استراتيجية البحث تتوقف على العوامل التالية:

1- قدرة القائم بالبحث على اختيار المصطلحات الصحيحة للتعبير عن موضوع البحث.
2- قدرة القائم بالبحث على وضع المصطلحات معا على نحو سليم منطقيا.
3- قدرة القائم بالبحث على تدبر جميع الطرق المناسبة للاسترجاع.
4- قدرة القائم بالبحث على صياغة الاستراتيجية بتعديل الشمول أو التخصيص بما يناسب متطلبات الاستدعاء والتحقيق أو قدرات احتمال المستفيد.

3/مستويات استراتيجية البحث:

يتوقف مستوى الاستراتيجية على طبيعة تساؤل المستفيد، وعلى العناصر التي يتشكل منها هذا التساؤل، ويوجد في هذا الصدد ثلاثة مستويات:

أ- استراتيجية بسيطة:

وهى التي تعكس تساؤلا مبسطا من جانب المستفيد، وتتطلب صياغتها مصطلحا واحدا دون الحاجة لربطه بمصطلحات اخرى، وعادة ما يتطلب تنفيذ هذه الاستراتيجية البحث في حقل واحد من الحقول القابلة للبحث في قاعدة البيانات، على سبيل المثال باحث في مجال المكتبات والمعلومات يرغب في التعرف على ما نشر عن موضوع" الدراسات الببليومترية" بصرف النظر عن أية حدود أخرى للموضوع.

ب-استراتيجية مركبة:

هي التي تعكس تساؤلا يتضمن جانبين موضوعيين، أو جانب موضوعي واحد يرتبط بجوانب أخرى شكلية أو لغوية أو جغرافية أو زمنية. فهب على سبيل المثال أن ذلك الباحث يرغب فى تحديد بحثه لموضوع الدراسات الببليومترية بحدود موضوعية أخر كالدراسات الببليومترية في مجال علم الاجتماع، أو بحدود لغوية، كأن يشترط أن تكون المواد المسترجعة منشورة باللغة الإنجليزية فقط، أو في مجلة علمية معينة. وفى هذه الحالة يتطلب الأمر صياغة الاستراتيجية بحيث

تتضمن الموضوع الأساسي مرتبطا بالجوانب الأخرى عن طريق الروابط المنطقية المناسبة.

ج- **استراتيجية معقدة:**

هي التي تعكس تساؤلا يتضمن عدة جوانب لكل منها أهميته بالنسبة للباحث، كأن يكون لهذا الاستفسار أكثر من جانب موضوعي بالإضافة إلى بعض المحددات اللغوية أو الزمنية أو الجغرافية، فعلي سبيل المثال يرغب ذلك الباحث في التعرف على ما كتب عن موضوع" الدراسات الببليومترية أو تحليل الاستشهات المرجعية في جميع مجالات العلوم الاجتماعية باستثناء علم الاجتماع خلال الفترة الزمنية من 1995- 1999، وبشرط أن تكون منشورة باللغة الانجليزية فقط"، فمن الواضح أن هذا الاستفسار يتضمن عدة جوانب موضوعية، ويتطلب استخدام عدة روابط منطقية، وبعض المحددات الحقلية.

4/ آلية تنفيذ استراتيجية البحث:

بعد الاستقرار على الصياغة الصحيحة لاستراتيجية البحث، تتم كتابتها في المكان المخصص لذلك بناء على نوع واجهة التعامل وشكلها، والأسلوب الذى تتيحه برمجية البحث المستخدمة، ويتم تنفيذ

الاستراتيجية بعد صياغتها بأحد أسلوبين يتم اختيار أحدهما استنادا إلى ثلاثة اعتبارات تتمثل في مستوى الاستراتيجية، وكفاءة القائم على تنفيذها، وإمكانات برمجية البحث والاسترجاع المستخدمة .

يتمثل الأسلوب الأول في الأسلوب المرحلي، أي أن تجزأ استراتيجية البحث مرحليا بعدد عناصر الاستفسار، على أن يتم البحث عن كل عنصر بشكل مستقل، ومن ثم ينتج عن ذلك استرجاع مجموعة الوثائق التي تناسب هذا العنصر، ثم يتم في النهاية دمج البحوث المرحلية السابقة في خطوة واحدة عن طريق أرقامها أو استخدام رموز خاصة تحددها البرمجية. أم الأسلوب الثاني فيتمثل في البحث المركب، وهو إجراء البحث في خطوة واحدة، بحيث تشمل هذه الخطوة جميع الجوانب التي يتضمنها التساؤل.

5/ تقنيات البحث:

إننا أمام رصيد متزايد من قواعد البيانات المتفاوتة، نتيجة لتعدد برمجيات البحث والاسترجاع وتباين مصادر إنتاجها، وتعدد إمكاناتها وخصائصها.

ونركز هنا على إحدى هذه الإمكانات، وهى الإمكانات المتصلة بتقنيات البحث، أي الإمكانات والأساليب التي توفرها البرمجيات بهدف ضبط عمليات البحث وتنظيمها من أجل الارتفاع بمعدلات التحقيق والحد من الاسترجاع الخاطئ، وتجنب الربط الزائف بين المصطلحات في استراتيجيات البحث، ويمكن أن تصنف تقنيات البحث إلى أربع فئات على النحو التالي:

أ- تقنيات البتر.

ب-تقنيات الربط بين المصطلحات.

ت-تقنيات تقييد البحث.

ث-سبل التعامل مع كلمات التوقف.

أ- تقنيات البتر Truncation:

البتر لغة هو القطع، واصطلاحا يعنى إخفاء أو حذف اللواصق المرتبطة بجذور الكلمات، سواء أكانت هذه اللواصق سوابق أم لواحق، والاستعاضة عن تلك اللواصق المحذوفة بحروف أو رموز معينة ترف برموز البتر truncation symbols. وتعد تقنية البتر من أساليب البحث الهامة التي ارتبطت أساسا بنظم الاسترجاع المعتمدة على اللغة الطبيعية، أو قواعد بيانات النصوص الكاملة، إلا أن استخدامها لم يعد

يقتصر على تلك القواعد، بل كادت تستخدم في جميع قواعد البيانات على اختلاف أنواعها. والهدف الأساسي لتقنية البتر يتمثل في توسعة مجال البحث بحيث يتم استرجاع جميع المصطلحات ذات العلاقة، حتى وإن كانت هذه العلاقة علاقة شكلية وليست موضوعية.

- أنواع البتر:

للبتر ثلاثة أنواع:

أ- البتر الأيمن

ويقصد به حذف اللواحق التي تأتى على يمكين جذر الكلمة، أو بعد مقطع معين من الكلمة، هذا بالنسبة للغات التي تكتب من اليسار إلى اليمين، في حين تأتى هذه اللواحق على اليسار في حالة اللغات التي تكتب من اليمين إلى اليسار. فكلمة computer عندما يراد أن تسترجع كل مشتقاتها تبتر بترا أيمنا على النحو التالي *comput في هذه الحالة تسترجع المشتقات التالية على سبيل المثال لا الحصر:

Compute, computer, computing, computerized, computation, computational

وتجدر الإشارة إلى أن تقنية البتر الأيمن لا تقتصر على الكلمات أو المصطلحات المفردة، وإنما يمكن أن يتم لجزء من مصطلح مركب، كالاسم الأول التالي لاسم العائلة بالنسبة لأسماء الأشخاص، كأن يكتب الاسم التالي هكذا: * ,sayed في هذه الحالة يتم استرجاع جميع أسماء المؤلفين الذين يشتركون في اسم العائلة sayed.

ب-البتر الأيسر:

يقصد به حذف السوابق التي تأتى علي يسار جذر الكلمة أو مقطع معين منها، هذا بالنسبة للغات التي تكتب من اليسار إلى اليمين، أما بالنسبة للغات التي تكتب من اليمين إلى اليسار فتأتى هذه السوابق على يمين جذر الكلمة لا على يساره. وعادة ما تستخدم هذه التقنية لاسترجاع المصطلحات ذات اللواحق المتشابهة والبوادئ المختلفة، وبخاصة فى قواعد البيانات المتخصصة فى المجالات العلمية كالطب والصيدلة بشكل عام، وذلك لاسترجاع المصطلحات العلمية أو المركبات الدوائية والكيميائية وما شابه ذلك. فعل سبيل المثال باحث يرغب في استرجاع المركبات المنهية بالمقطع"pirin"، في هذه الحالية يستخدم تقنية البتر الأيسر في كتابة المصطلح على النحو التالي"pirin*" ومن ثم تسترجع المصطلحات التالية:

Algopirin, Aspirin, Kalmopirin

ويعد هذا النوع أصعب في تنفيذه من النوع الأول بالنسبة للغات التي تكتب من اليسار إلى اليمين، ومن ثم فإن فئة معينة فقط من البرمجيات هي التي تتيح هذا النوع من البتر وبحدود معينة.

ت-البتر الداخلي أو الأوسط:

ويقصد به حذف أو إخفاء حرف أو أكثر من وسط الكلمة أو أصلها وليس من لواحقها أو سوابقها. وعادة ما يستخدم هذا الأسلوب فى الحالات الثلاثة التالية:

1- التفاوت الشكلي غير الملحوظ بين الهجائية الأمريكية والهجائية البريطانية في بعض الكلمات، ومن أمثلة ذلك:

Color-colour

Catalogue-cataloge

2- تفاوت كتابة بعض الأسماء والمصطلحات عند نقلها نقلا حرفيا أو صوتيا من لغتها الأصلية إلى لغة أخرى من أصل مختلف،

فالاسم "محمد: تتفاوت صيغ كتابته بالحروف اللاتينية على النحو التالي

Mohamad, Mohammad, Mohamed, Muhamad,...

ومن ثم فإن كتابة الاسم بصيغة غير المخزنة فى النظام يترتب عليها عدم الاسترجاع.

3- الجمع غير القياسي، فالجمع القياسي في اللغة الانجليزية عادة ما يتم عن طريق إضافة حرف sأو es أو ss لنهايات الكلمات. لكن هناك الكثير من الجموع غير القياسية أو الشاذة لبعض الكلمات، وبالتالي ينبغى تحرى الدقة عند البحث عن مثل هذه الكلمات، وما لم تكتب بصيغتها الصحيحة فلا يتم استرجاعها. ومن أمثلة تلك الكلمات:

Woman- Women

Feet- Foot

Thesis- theses

- **رموز البتر:**

عادة ما يستعاض عن الحروف المبتورة أو المقنعة برموز معينة تعرف" برموز البتر، والاختلافات واضحة بين برمجيات البحث والاسترجاع فى شكل هذه الرموز؛ فهناك برمجيات تستخدم رمزا واحدا للبتر بجميع أنواعه، وهناك برمجيات أخرى تخصص رمزا لكل من البتر الأيمن والأيسر، ورمزا آخر للبتر الداخلى، وهناك برمجيات تخصص رمزا لبتر حرف واحد، ورمزا ىخر لبتر أكثر من حرف

ومن الرموز الشائعة الاستخدام في هذا الصدد:

- النجمة(*): وتستخدم من قبل برمجيات مثل: EBSCO, ProQuest, Cambridge
- علام الاستفهام(?): وتستخدم من قبل برمجيات مثل Dialogo on Disc, Bowker:
- الشارحة(:): وتستخدم من قبل برمجيات مثل Wilsonline, Wilsearch:
- علامة (#): وتستخدم كذلك من قبل برمجيات: Wilsonline, Wilsearch

- علامة الدولار ($): وتنفرد باستخدامها برمجية Bowker- CD فى حالة البتر الداخلي.

وتجدر الإشارة إلى أن معظم البرمجيات تستخدم رمز البتر مرة واحدة للدلالة على بتر حرف واحد أو أكثر، فى حين تلجأ بعض البرمجيات إلى تكرار رمز البتر بعدد الحروف المبتورة.

ب-تقنية الربط بين المصطلحات:

أشرنا سابقا أن اختيار مصطلحات البحث عادة ما يتم بناء على طبيعة الاستفسار، فقد يكون استفسارا بسيطا، وبالتالي يكتفى فى التعبير عنه بمصطلح واحد، وقد يكون استفسارا مركبا أو معقدا، وبالتالي يكتفى في التعبير عنه بمصطلح واحد، وقد يكون استفسارا مركبا أو معقدا، وبالتالي يتطلب مجموعة من المصطلحات التي يلبى كل منها جانبا من جوانب الاستفسار، وفى هذه الحالة يتطلب الموقف أن تربط هذه المصطلحات فيما بينها بروابط تعرف بالروابط المنطقيةlogical operators، بحيث تشكل في مجموعها التعبير الدقيق عن حاجة المستفيد.

ويمكن تقسيم هذه الروابط المنطقية إلى عدة أنواع بناء على طبيعة العلاقة التي يرغب الباحث في بنائها بين مصطلحات البحث، حيث يمكن حصر الروابط التي تستخدم في بناء العلاقات المنطقية بين جوانب الاستفسارات، وبالتالي تطبق في صياغة استراتيجيات البحث، في الأنواع الأربعة التالية:

- الروابط البولينية.
- الروابط التجاورية أو التقاربية.
- الروابط الموضعية.

1- الروابط البولينية Boolean operators:

وهى الروابط التي تستخدم للربط الموضوعي بين مصطلحات البحث، بهدف بناء علاقة عضوية بين تلك المصطلحات. وتنسب هذه الروابط إلى عالم الرياضيات والمنطق البريطانى جورج بول، وتتمثل هذه الروابط فى الكلمات الثلاث التالية AND, OR, NOT

1/1 الرابط AND(و):

وتسمى العلاقة التي يعبر عنها بهذا الرابط علاقة الاقتران، أي تلازم جميع الجوانب التي تعبر عن حاجة المستفيد في الوثائق

المسترجعة تلازما عضويا لا انفكاك فيه، فعلى سبيل المثال باحث يرغب في التعرف على الوثائق التي تتناول موضوع" المكتبات العامة في العالم العربي"، فهذا التساؤل يتشكل من جانبين هما: "المكتبات العامة" و" العالم العربي" ومن ثم فإن الوثائق المسترجعة التي تلبى الحاجة الفعلية للمستفيد ينبغي أن تتناول هذين الجانبين بالتحديد معا، ودون سواهما من موضوعات. فالوثائق التي تتحدث عن المكتبات بشكل عام، أو التي تتحدث عن المكتبات العامة في مناطق أخرى غير العالم العربي لا تلبى حاجة المستفيد، كما أن الوثائق التي تتناول العالم العربي كمنطقة جغرافية لا تلبى حاجة المستفيد كذلك. ومن ثم يجب أن تصاغ استراتيجية البحث على النحو التالي:

Public Libraries AND Arab world

2/1 الرابط OR(أو):

في حالات كثيرة يتضمن تساؤل المستفيد عدة جوانب بينها علاقة بدل، وعادة ما تكون بين طرفين متجانسين؛ حيث يمكن لأحدهما أن يحل محل الآخر، وتتعدد أوجه العلاقات الى عادة ما

تتضمنها تساؤلات الباحثين وتتطلب استخدام علاقة البدل OR والتي تعد من قبيل الاستراتيجيات المعقدة أو المركبة، ومن أهم تلك الأوجه ما يلى:

أ- جانب موضوعي واحد لا يكفى لضمان استرجاع ما يتصل به، التعبير عنه بمصطلح واحد فى استراتيجية البحث وإنما بأكثر من مصطلح، نتيجة لتعدد المترادفات التي تعبر عنه، فعلى سبيل المثال مصطلح " الدراسات الببليومترية" يعبر عنه أيضا ب " الببليوجرافية الاحصائية أو القياسات الوراقية". ومن ثم فالبحث بمصطلح واحد دون الآخر له تأثير سلبى على نتائج البحث. بحيث تصاغ استراتيجية البحث هكذا

Bibliometrics or Statistical Bibliography

ب- جانب موضوعي ذو بعد جغرافي ولكن يعبر عنه بأكثر من مرادف، كما في مثال المكتبات العامة في العالم العربي، حيث قد يعبر عن البعد الجغرافي بمصطلح الدول العربية. ومن ثم تصاغ استراتيجية البحث على النحو التالي:

Public libraries AND (Arab world OR Arab countries)

ث- موضوع عام يتضمن عدة جوانب فرعية، وكثيرا ما تستخدم المصطلحات الدالة على هذه الجوانب الفرعية بشكل يكاد يكون تبادليا، فعلى سبيل المثال "الدراسات الببليومترية" أوسع وأشمل من " تحليل الاستشهادات المرجعية" ولكن يستخدم المصطلح الأخير في كثير من الاحيان للتعبير عن الأول.

فتصاغ استراتيجية البحث على النحو التاليى:

Citation Analysis OR Bibliometrics

3/1 الرابط NOT(فيما عدا):

تعرف العلاقة التي تستخدم هذا الرابط في التعبير عنها بعلاقة الاستبعاد، ففي المثال السابق قد يرغب الباحث في استبعاد دولة معينة، كأن يحدد بحثه بما كتب عن "المكتبات العامة في العالم العربي باستثناء مصر" ومن ثم تصاغ الاستراتيجية على النحو التالي:

Public Libraries AND (Arab world NOT Egypt)

وهذ الصياغة تعنى استرجاع الوثائق التي تتناول:

- المكتبات العامة في العالم العربي فيما عدا مصر.
- أو المكتبات العامة في العالم العربي وليس مصر فقط.
- أو المكتبات العامة في العالم العربي فقط، وليس العالم العربي ومصر معا.

2-الروابط التقاربية أو التجاورية:

ارتبطت تقنية البحث بالتجاور أو بالتقارب بقواعد بيانات النصوص الكاملة؛ حيث تعد أحد أساليب بحث النصوص الطبيعية او الحرة، ثم اتسع مجال استخدامها في قواعد البيانات الببليوجرافية التي تشتمل على مستخلصات إعلامية مطولة. والهدف الأساسي من تطبيق تقنية البحث بالتجاور هو الحد قدر الإمكان من الاسترجاع الخاطئ نتيجة الربط الزائف للمصطلحات والذى يحدث عند استخدام الروابط المنطقية، وبخاصة الرابط AND ، حيث يمكن لاستراتيجية بحث مصاغة على النحو التالي" Public Libraries AND Egypt" أن يتم بموجبها استرجاع جميع النصوص التي ورد فيها المصطلحان بصرف النظر عما إذا كانت هناك علاقة موضوعية قوية لهذه النصوص بموضوع الاستفسار أم لا؛ فقد تسترجع وثيقة ورد بهها مصطلح "مصر" في مكان النشر

الكتاب الذى ورد في عنوانه مصطلح" المكتبات العامة". إلى أخر ذلك من احتمالات الاسترجاع الخاطئ.

ومن هنا كان التفكير في إتاحة هذه التقنية حيث يمكن للقائم بصياغة استراتيجية البحث أن يحدد طول المسافة الفاصلة بين المصطلحات التي يتم الربط بينها، وعادة ما يتم تحديد طول هذه المسافة بعدد الكلمات الفاصلة بين المصطلحات الاساسية. وتستند فكرة هذه التقنية إلى فرضية بانه كلما تقاربت المصطلحات المرتبطة، قويت العلاقة الموضوعية بينها.

وهناك عدة مستويات للربط بين المصطلحات ربطا تجاوريا، أهمها المستويات التالية:

- الربط بين المصطلحات على مستوى الحقل.
- الربط بين المصطلحات على مستوى الفقرة.
- الربط بين المصطلحات على مستوى الجملة.
- الربط بين المصطلحات على مستوى الكلمة.
- الربط بين المصطلحات على مستوى الواصفات الموضوعية.

ويعنى الربط بين المصطلحات على مستوى الحقل، أن تلك المصطلحات التي تم تحديدها في استراتيجية البحث يجب أن تتواجد معا داخل نفس الحقل الموضوعي الذى تم تحديده للبحث، مما يقلل من الربط الزائف بين المصطلحات.

وتتفاوت النظم التي تتيح برمجياتها خاصية البحث بالتقارب تفاوتا واضحا في أسلوب التعبير عن تلك العلاقة وفى نوع الرابط الذى يستخدم.

3-الروابط الموضعية:

في تطبيق تقنية البحث بالتقارب لم تكن هناك ضرورة لترتيب مصطلحات البحث، لأن ذلك لم يكن له تأثير على طبيعة التساؤل. ولكن في بعض الاحيان وبخاصة عند البحث في قواعد بيانات النصوص الكاملة يكون ترتيب المصطلحات ضروريا، وما لم ترتب المصطلحات فإن النتائج ستكون غير ملبية للحاجة الحقيقية للمستفيد، فلا شك أن الفارق واضح بين المصطلحات التالية:

- "إدارة الأعمال" و "أعمال الإدارة".
- "كليات المكتبات" و" مكتبات الكليات".

ومن هنا كان التفكير في تقنية الربط الموضعي، وذلك عن طريق تحديد ترتيب المصطلحات في استراتيجية البحث ترتيبا طبيعيا أو دلاليا يضمن تلازم مصطلحات البحث في حدود فاصل معين، وبترتيب معين، هذا وتتفاوت البرمجيات التي تتيح تقنية الربط الموضعي في أشكال الروابط أو الرموز المستخدمة للتعبير عن تلك العلاقة.

ت-تقنية التقييد الحقلي:

أشرنا سابقا إلى أن برمجيات البحث والاسترجاع تحرص قدر الاستطاعة على الحد من الاسترجاع الخاطئ للربط المزيف بين المصطلحات، ومن ثم كانت تقنيات البحث بالتقارب، وتقنيات الربط الموضعي، ومع ذلك لم تكن تلك الروابط الحل الأمثل في جميع الأحوال، هذ فضلا عن أن الكثير من البرمجيات لم تتح هذه التقنيات.

وفضلا عن تقنيات البحث بالتقارب والربط الموضعي، تتيح معظم البرمجيات تقنية "التقييد الحقلي "أي تقييد الحقل الذى يتم البحث فيه عن المصطلحات الأساسية التي تم تحديدها في استراتيجية البحث. وما لم يتم تحديد حقل معين يفترض النظام أن الباحث يرغب أن يتم البحث

حرا في جميع الحقول، وبالتالي يتم استرجاع التسجيلات التي وردت فيها مصطلحات البحث بصرف النظر عن مكان ورودها، ففي المثال السابق " المكتبات العامة في مصر" ما لم يتم تحديد الحقل الذى يتم البحث فيه عن المصطلحات الأساسية للبحث وهى "المكتبات العامة" و"مصر" سوف يتم استرجاع جميع الوثائق التي ورد هذان المصطلحان في سياق عناوينها، أو في حقل الموضوع، أو الواصفات، أو المستخلصات، أو حقل النشر، أو غيرها من الحقول، بصرف النظر عن وجود علاقة بين الوثائق المسترجعة وموضوع التساؤل. أما إذا تم تحديد البحث عن المصطلحات الأساسية في حقل معين كحقل العنوان مثلا فإن هذا بمثابة ضمان لاسترجاع الوثائق في سياق عناوينها تلك المصطلحات متلازمة.

وتتاح تقنية التقييد الحقلي من خلال إتاحة جميع الحقول القابلة للبحث كبدائل أو خيارات بحيث يقوم الباحث بتحديد الخيار المناسب كأن يكون حقل العنوان، وذلك بعد كتابة مصطلحات البحث في المكان المخصص كما هو موضح في الشكل رقم (1).

شكل رقم (1) شاشة بحث بأسلوب التقييد الحقلى فى قاعدة Education Plus العاملة ببرمجية ProQuest

ث- كلمات التوقف وسبل التعامل معهاstopwords :

يقصد بكلمات التوقف، أو الكلمات المتخطاة كما يفضل البعض تسميتها، هي الكلمات غير ذات الدلالة اللغوية عند استخدامها مستقلة عن سياق الجملة أو العبارة، وبالتالي تعد هذه الكلمات غير قابلة للبحث، وينبغي ألا تستخدم ككلمات مفتاحية في استراتيجيات البحث. وتشمل قائمة الكلمات المستبعدة عدة فئات، من أهمها ما يلى:

- حروف الجر والعطف والضمائر وأسماء الإشارة والأسماء الموصولة وأدوات التعريف وأدوات الشرط وأدوات الاستفهام...إلخ
- أدوات الربط المنطقى بأشكاله المختلفة.
- محددات الحقول القابلة للبحث أو رموزها.
- علامات الترقيم والرموز الخاصة المستخدمة للدلالة على بعض تقنيات البحث، مثل عمليات البتر والاختصار والتقييد...إلخ.

وتتفاوت قواعد البيانات في حجم قوائم كلمات التوقف بها . إلا أن يمكننا القول بأن جميع قواعد البيانات تحرص على إيراد قائمة بالكلمات غير القابلة للبحث في دليلها الإرشادي.

ولا شك أن موقف كل من الباحث والبرمجية مع كلمات التوقف له احتمالين لا ثالث لهما/ موقف يتطلب صياغة استراتيجية بحث معينة، ولكن ترد في سياقها إحدى كلمات التوقف على سبيل الخطأ أو عدم الوعى الكافي بالأسس الصحيحة لصياغة استراتيجية البحث، وبالتالي فإن حذف مثل هذه الكلمات من قبل النظام ليس له أي تأثير على عملية البحث أو على معدلات التحقيق، ويحدث

هذا الموقف في حالة البحوث المبسطة في قواعد البيانات الببليوجرافية.

أما الموقف الثاني فهو يتطلب حتما ان ترد كلمة واحدة أو اكثر من كلمات التوقف في سياق استراتيجية البحث التي يفرضها هذا الموقف، وبالتالي لا يمكن ان تحذف هذه الكلمات من جانب النظام أو أن يتم تجاهلها من جانب الباحث لما لذلك من تأثير على معدلات التحقيق، وهذا الموقف يظهر في حالة بحث الجمل والعبارات سواء في قواعد البيانات النصوص الكاملة أو قواعد البيانات الببليوجرافية المشتملة على مستخلصات مطولة.

وإذا لم يلتزم الباحث عند صياغة استراتيجية البحث بتعليمات البرمجية وسبل تعاملها معل كلمات التوقف، فإن نتائج البحث تكون سلبية.

وقد بدا لنا عبر النقاط التي تناولناها أن هناك قدرا كبيرا من التفاوت بين برمجيات البحث في تقنياتها وإمكانياتها البحثية، وهذا التفاوت مرده في المقام الأول يرجع لعاملين أساسيين، يتمثل أولهما في التزايد المطرد لإنتاج قواعد البيانات وحدة روح التنافس بين تلك

الشركات المنتجة، أما العامل الآخر فيمثل في غياب جهود التقييس وتوحيد المواصفات أو بالأحرى ضعفها في هذا المجال، وإذا لم يكن لنا يد في كل سلبيات هذا التفاوت الواضح، فلا سبيل أمامنا سوى التأقلم مع هذه السلبيات وذلك بالحرص الشديد على الاسترشاد بالنشرات التوثيقية والموجزات الإرشادية التي تصاحب كل قاعدة من قواعد البيانات.

قائمة المصادر والمراجع

أولا: المصادر والمراجع العربية:

أحمد أنور بدر (يوليو 2009) تطور الإنتاج الفكري لمجال استرجاع المعلومات وواقع تدريسه في بعض أقسام المكتبات والمعلومات المصرية والسعودية: دراسة توثيقية ميدانية. الاتجاهات الحديثة في المكتبات والمعلومات، مج 16، ع 32 ص11 – 54.

أحمد أنور بدر (2000) أساسيات استرجاع المعلومات: ص57 – 78 في كتابه تكنولوجيا المعلومات وأساسيات استرجاع المعلومات – الإسكندرية: دار الثقافة العلمية، 100 ص.

أحمد أنور بدر ومحمد فتحي عبد الهادي ونايمان متولي (2001) التكشيف والاستخلاص: دراسات في التحليل الموضوعي – القاهرة، دار قباء، 416 ص.

أحمد أنور بدر، زهير سليمان مالكي ونايمان إسماعيل متولي (2013) مقدمة في علم المعلومات: النشأة والتطور حتى العصر الرقمي – القاهرة: المكتبة الأكاديمية، 296 ص.

حشمت قاسم (1985) كشافات الكلمات المفتاحية في السياق واحتمالاته في اللغة العربية عالم الكتب، مج 5، ع4، ص638 – 650.

داود عطية عبده (1979) المفردات الشائعة في اللغة العربية. الرياض: جامعة الرياض، معهد اللغة العربية، 343ص.

- علي سليمان الصوينع (1994) استرجاع المعلومات في اللغة العربية. الرياض: مكتبة الملك فهد الوطنية، 176 ص.
- فتحي عثمان أبو النجا (1985) وضع نظام عربي لاسترجاع المعلومات في قطاع المعلومات الزراعية (دكتوراه جامعة القاهرة).
- كنت، ألن (1971) ثورة المعلومات: استخدام الحاسبات الإلكترونية في اختزان المعلومات واسترجاعها: ترجمة حشمت قاسم وشوقي سالم، مراجعة أحمد بدر، الكويت، وكالة المطبوعات، 984 ص.
- محمد سالم غنيم (2008) نظم استرجاع المعلومات العربية: مظاهر الغموض وآفاق الحلول: الرياض: مكتبة الملك فهد الوطنية، 474 ص (أطروحة دكتوراه) القاهرة.
- محمد بن عبد الله الأطرم (1990) كفاءة اللغة العربية في تكشيف واسترجاع الوثائق العربية التقرير النهائي الرياض: مدينة الملك عبد العزيز للعلوم والتقنية.
- مساعد بن صالح الطيار (1998) كفاءة التحليل الصرفي في استرجاع النصوص العربية – مجلة مكتبة الملك فهد الوطنية – مج4، ع1، ص7 – 23.
- ناريمان إسماعيل متولي (يناير 2004) الإنترنت والأطر البحثية في استرجاع المعلومات. الاتجاهات الحديثة في المكتبات والمعلومات – ع21 – ص55 – 77.

- نبيل عبد الرحمن المعثم (نوفمبر 2011) البحث باللغة العربية على محرك البحث جوجل – مجلة مكتبة الملك فهد الوطنية – مج7، ع2.
- نزار محمد علي قاسم (2001) إمكانية رفع كفاية استرجاع المعلومات باستخدام خصائص المفردات العربية (أطروحة دكتوراه) الجامعة المستنصرية، قسم المكتبات والمعلومات ص180.
- نبيل علي (1988) اللغة العربية والحاسوب: دراسة بحثية تقديم أسامة الخولي – القاهرة تعريب للنشر، 591 ص.

ثانيا: المصادر والمراجع الأجنبية:

- Abdelali, Ahmed; Soliman, Hand}/ (2004) Arab Information Retrieval, Arab Language Processing, v.19.
- Abu El Khair (2006) Effects of stop Words Elimination for AIR: Comparative stud)/ international). of Computing and Information Science v.4 (3) On-line 119-132.
- Abu El Khair, I. (2003). Effectiveness of document processing techniques for Arabic informationretrieval. Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh.
- Abu El Khair (2007) Information Retrieval ARIST, v.41 (1) ch.11 p505-531.

- Al Dayal, Abeer and Mourad jykh/ef (2013). Arabic user's attitudes toward web searching using paraphrasing mechanisms journal of Compute Science and information systems, v.2(2), p 34-39.
- Aljlayl, Mohamed and Frieden, Ophir (2002) On Arabic Search: improving the Retrieval effectiveness via a light stemming Approach CIKM 02 November 4-9, Mclean, Virginia, U.S.A.
- Al-Shehri, Abd-Allah M. (2002). Optimization and effectiveness of N-grams approach for indexing and retrieval in Arabic information retrieval systems. Thesis(ph.D.) university of Pittsburgh, A. Al Shehri, 2002173p. Pittsburgh.
- Al Tayyar, Musaid Saleh (july 2000). Arabic information Retrieval System based on Morphological Analysis: A comparative study of word, stem, root and morph semantic Methods. Ph.D. in computer science in the department of computer and information science, De Mont fort university, U.S.A.
- Attia, M. (2006). An Ambiguity-Controlled Morphological Analyzer for Modern Standard Arabic Modeling Finite State Networks. In Proceedings of the Challenge of Arabic for NLP/MT Conference. The British Computer Society, London.

- Chen, Atitao and Gey, F. (2002) Building an Arabic stemmer for information Retrieval In: Proceeding of TREC 2000.
- Chowdhury, G. S. (2004). Introduction to Modern Retrieval. Facet Publishing, London.
- Darwish, K. (2002a). Al-stem: A light Arabic Stemmer. Retrieval june 20, 2005, from www.glue.umd.edu/~kareem/research.
- Darwish, K. (2002b, July). Building a shallow Arabic morphological analyzer in one day. Proceeding of the ACL-02 Workshop on Computational Approaches to Semitic Languages, University of Pennsylvania, Pennsylvania, PA, USA. Retrieval January 26, 2006, from the http://acl.ldc.upenn.edu/w/w02-0506.pdf..
- Darwish, K, & Oard, D,, (2002). CLIR experiments at Maryland for TREC-2002: Evidence combination for Arabic-English retrieval. The Eleventh Text Retrieval Conference (TREC 2002), 703-710. Retrieved January 25, 2006, from http://trec.nist.gov/pubs/trec11/papers/umd.darwish.pdf.
- Harman, Donna et al (2004). The NRRC reliable information access (RLA) workshop (SIGIR) pp528-529..

- Hersh, William et al (2006) TREC: Genomics Track Overview In: Woorkees & Buchlard (2006.
- Kadri, Y. and Nie,). Y. (2006). Effective Stemming for Arabic Information Retrieval, the Challenge of Arabic for NLP/MT. In Proceedings of the International Conference ar the British Computer Scociety, pp68-74, London, UK.
- Kando, Noriko and Evans, David. (2007). Proceedings of the 6th NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering, and Cross-Lingual Information Access. National Institute of Informatics, Tokoyo,)apan.
- Khoja, S. (2001). Khoja's Arabic stemmer (version 1.0). London: Khoja.
- Khoja, S. & Garside, R. (1999). Stemming Arabic text. Retrieval from www.comp.lancs.ac.uk/computing/users/khoia/stemmer.ps
- otair, Mohammed et al (2013). Optimizing an Arabic Query using Comprehensive Query Expansion Techniques. International journal of Computer Applications, v. 71 (17), 42-49