

جامعة السودان للعلوم والتكنولوجيا  
كلية علوم الحاسوب وتقانة المعلومات  
قسم الحاسوب ونظم المعلومات

# بناء مدونة قياسية باللغة العربية

## Creating Standard Corpus For Arabic Language

مشروع م قدم كأحد متطلبات الحصول على بكالوريوس الشرف  
في الحاسوب ونظم المعلومات

اكتوبر 2016



بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

جامعة السودان للعلوم والتكنولوجيا

كلية علوم الحاسوب و تقانة المعلومات

قسم الحاسوب ونظم المعلومات

بناء مدونة قياسية باللغة العربية

## Creating Standard Corpus

### For Arabic Language

إعداد :

• أسماء عبدالوهاب نوار

• كوثر محمد آدم

• منيره متوكل ماريو

التاريخ:

اكتوبر

الإشراف:

إبتهال مصطفى

2016

# الآية

قال تعالى:

وَمَا أُوتِيتُمْ مِنَ الْعِلْمِ إِلَّا قَلِيلًا)

صدق الله العظيم

سورة الإسراء الآية (85)

# الحمْد

الحمْد والشكر لله الذي وفقنا وأعاننا على إكمال هذا العمل المتواضع ونخص بالشكر والتقدير الأستاذة إبتهاال مصطفى المشرفة علي هذا البحث لتوجيهاتها وإرشاداتها التي كانت لها اثر كبير في إخراج هذا البحث.

ونخص أيضا بالشكر الأستاذ محمد عدني الذي اجتهد معنا وكان نعم الموجه والمرشد وفقه الله لخدمة العلم وطلابه.

ولا يسعنا في الختام إلا أن نتقدم بشكرنا وتقديرنا لكل من ساهم وساعد في إخراج هذا البحث المتواضع.

# الإهداء

أهدي هذا العمل إلى

من ربّنتني وأنارت دربي وأعانتني بالصلوات والدعوات، إلى أعلى إنسان في هذا الوجود

أمي الحبيبة

إلى من عمل بكدي في سبيلي وعلمني معنى الكفاح وأوصلني إلى ما أنا عليه

أبي الكريم أدامه الله لي

إلى من حبهم يجري في عروقي ويلهج بذكراهم فؤادي.. وودعموني شتّى أنواع الدعم...

إخوتي وأخواتي

إلى من عمل معي بكدي بغية إتمام هذا العمل

أصدقائي

وفي الأخير أرجو من الله تعالى أن يجعل هذا العمل نافعا يستفيد منه جميع الطلاب المقبلين على

التخرّج

## الشكر والعرفان

أشكر الله العليّ القدير الذي أنعم عليّ بنعمة العقل والدين. القائل في محكم التنزيل وَفَوْقَ كُلِّ ذِي  
لُحْمٍ عَلِيمٌ

سورة يوسف (آية 76) صدق الله العظيم .

و قال رسول الله صلي الله عليه وسلم): "من صنع إليكم معروفاً فكافئوه, فإن لم تجدوا ما تكافئونه به فادعوا له حتى تروا أنكم كافأتموه" ( رواه أبو داوود" .

وفاءً وتقديرًا وإعترافًا منا بالجميل نتقدم بجزيل الشكر لأولئك المخلصين الذين لم يألوا جهداً في مساعدتنا في مجال البحث العلمي، ونخص بالذكر الأستاذة الفاضلة : إبتهاال مصطفى على هذه الدراسة وصاحبة الفضل في توجيهي ومساعدتي في تجميع المادة البحثية، فجزاها الله كل خير. وأخيراً، نتقدم بجزيل شكرنا إلي كل من مدوا لي يد العون والمساعدة في إخراج هذه الدراسة على أكمل وجه.

## المستخلص

يهدف البحث الى إسترجاع البيانات المخزنة وفقا لإحتياجات المستخدم عن طريق توفير مدونة قياسية باللغة العربية وهي عبارة عن مجموعة من المواضيع متعلقة فقط بمجال الحاسوب .

المدونة تحتوي على مجموعة من المستندات و مجموعة من الإستفسارات والحكم الملائم ( Relevance Judgment), يستطيع أي مستخدم الحصول على المدونة وإستخدامها في عملية الإسترجاع .

تم بناء المدونة بواسطة جمع بيانات متعلقة بمجال الحاسوب وحفظها وفهرستها بطريقة معينة ومن ثم إجراء عدة عمليات إحصائية عليها بواسطة برنامج (rapidminer) التي تمكن من معرفة عدد الكلمات الموجودة في المستندات و عدد الكلمات الفريدة, وبعد ذلك يتم فهرسة المستندات والبحث عن المعلومات بإدخال الإستفسار بواسطة برنامج (lucene), وتكون نتيجة إدخال الإستفسار هي إسترجاع جميع المستندات ذات الصلة بالإستفسار .

# Abstract

The project work on retrieval of stored data system according to user needs, This research aims to provide a standard corpus in Arabic, which is about a range of topics related to all the computing field.

This corpus contain to set of documents, set of queries and relevance judgment any users enables to access the corpus and used in the process of retrieval.

The project was accomplished by collecting a set of data on the computing field

and saved and indexed in appropriate way and then perform several operations by

statistical rapidminer program to account all the words in the documents and

account Total Occurences, and lucene are indexing and searching the data inside the

documents.

As a result, when you enter the queries is retrieval of all relevant documents Related to the queries.

## قائمة المصطلحات

المصطلح	المعنى	الاختصار
Information Retrieval	إسترجاع المعلومات	IR
Documents	المستندات	
Query	الإستفسار	
Data	البيانات	
Indexing	الفهرسة	
Matching	المطابقة	
Ranking	الترتيب	
corpus	مدونة	
Relevance Judgment	الحكم الملائم	
Data base	قواعد البيانات	DB
Unstructured	غير مهيكلي	
Structure	مهيكلي	
Relevant	مناسب	

Irrelevant	غير مناسب	
Rate Precision	معدل الدقة	
Rate Recall	معدل الاستدعاء	
Cross Language Information Retrieval	لغات أسترجاع البيانات	CLIR
Stopword	الكلمات المهملة	
Tokenization	التجزئة	
Normalization and Stemming	تبسيط وتجريد الكلمة لجذرها	
Boolean Search Model	نموذج البحث البولياني	
Exact Match	التطابق التام	
Ranked Retrieval Models	نموذج الترتيب	
The Vector space Processing Model	نموذج حيز المتجهات	
Probabilistic Retrieval Model	نموذج الإسترجاع الإحتمالي	
Lucene	برنامج اللوسين	
Rapidminer	رييد ماينر	
Microsoft Excel	مايكروسوفت اكسل	
Apache	الاباتشي	
Java Development Kit	أدوات تطوير الجافا	JDK
Notepad ++	محرر نصوص لكتابة الاكواد	

EXtensible Markup  
Language

لغة الترميز القابلة للتوسيع

XML

## فهرس الأشكال

## فهرس الجداول

## فهرس المحتويات

1.1 المقدمة.....	8
1.2 إسترجاع المعلومات.....	8
1.3 مشكلة البحث :	8
1.4 أهمية البحث :	8
1.5 أهداف البحث:	8
1.6 حدود البحث:	8
1.7 منهجية البحث:	8
1.8 هيكلية البحث:	8
1.9 خطة البحث:	8
الجدول(1-1) يوضح خطة البحث.....	8
2 الإطار النظري.....	11
2.1 مقدمة.....	11
2.2 خصائص اللغة العربية:	11
2.3 نظام إسترجاع المعلومات :	11
2.4 العمليات الأساسية في إسترجاع المعلومات:	11
2.5 كيفية تقييم نظام إسترجاع المعلومات:	12
2.5.1 معدل الدقة (Precision Rate):.....	12

2.5.2 معدل الإستدعاء ( Rate Recall ) .....	12
2.6 نماذج إسترجاع المعلومات: .....	13
2.6.1 (Boolean Search Model): نموذج البحث البولياني .....	13
2.6.2 (Ranked Retrieval Models): نماذج الترتيب .....	13
2.7 معالجة النصوص في لغات إسترجاع المعلومات: .....	14
(Text Processing In Language Information Retrieval ( CLIR)) .....	14
2.7.1 التجزئة (Tokenization) : .....	14
2.7.2 الكلمات المهملة (Stopword) .....	14
2.7.3 التطبيع (Normalization): .....	14
2.7.4 التجريد (Stemming ) : .....	14
2.8 الدراسات السابقة: .....	14
2.8.1 الدراسة الأولى: .....	14
2.8.2 الدراسة الثانية: .....	14
2.8.3 الدراسة الثالثة: .....	14
3 الأدوات والتقنيات المستخدمة: .....	16
3.1 ربيد ماينر (Rapidminer) .....	16
3.2 الجافا (Java Development Kit (JDK: .....	17
3.3 Notepad++ .....	17
3.4 برنامج (Lucene): .....	17
4 الباب الرابع .....	20
4.1 المقدمة: .....	20
4.2 المدونة (Corpora) : .....	20
4.3 مجموعة المستندات: .....	20
4.4 مجموعة الإستفسار: .....	20
4.4.1 كيف تم بناء المدونة: .....	20
الباب الخامس .....	21
5 النتائج والتوصيات: .....	22
5.1 النتائج: .....	22
5.2 التوصيات: .....	22
5.3 الخاتمة: .....	22

# الباب الأول

الإطار العام

## 1.1 المقدمة

بفضل التكنولوجيا والتطورات الحديثة ظهرت في العديد من المجالات حقائق ومعارف جديدة كانت مجهولة، ومن ناحية المجالات العلمية تطورت العلوم وأصبحت لها مفاهيم وتقنيات وطرق حديثة، أدى ذلك التطور إلى زيادة البيانات وتراكمها ومع كثرة البيانات أصبح من الصعب الحصول على المعلومة المطلوبة لأن ذلك يتطلب جهد ووقت وبناءاً على ذلك كان لابد من طريقة لترتيب البيانات لتتمكن من الوصول إلى المعلومات بسهولة.

ظهرت العديد من التقنيات والأدوات التي تمكن من الوصول إلى المعلومة المطلوبة ولكل تقنية مميزاتا وعيوبها، ومن أهم الأدوات التي ساعدت في ترتيب البيانات وحفظها هو الحاسوب و يمتاز بخصائص مثل المعالجة والتخزين والإسترجاع ويساعد في الأداء بصورة أفضل.

لتم عملية إسترجاع المعلومات لابد من ترتيب وتخزين مجموعة من البيانات بطريقة معينة حيث يكون لكل مجموعة بيانات خصائص ومعايير وأهداف تم جمعها على أساسها وهذه المجموعة من البيانات نطلق عليها مفهوم المدونة، وبعد ذلك إيجاد أفضل الطرق التي تمكن من الوصول إلى المعلومة بسهولة وبأقل وقت.

## 1.2 إسترجاع المعلومات

علم إسترجاع المعلومات من العلوم الهامة، وهو يقوم على عدة علوم من أهمها علم المكتبات، في السابق كان يعتمد على نظم إسترجاع المعلومات في المكتبات، فإذا كانت المكتبة تحتوي على مجموعة كبيرة من الكتب فيجب فهرسة هذه الكتب بطريقة معينة، لترتيب وتسهيل الوصول إليها، ومع تطور العلوم وتقدمها ظهرت الحاسبات وتم إستخدامها في أنظمة إسترجاع المعلومات الإلكترونية. [3]

توجد عدة أنواع لأنظمة إسترجاع المعلومات وسوف يركز البحث على إسترجاع المعلومات من المستندات.

يمكن تعريف علم إسترجاع المعلومات بأنه علم يبحث عن المعلومات داخل مجموعة كبيرة من المستندات (Documents) غير المهيكلة، وفقاً لطلب المستخدم، وهذه المستندات يتم جمعها وتخزينها بطريقة معينة و تتم فهرستها للوصول إليها بسهولة، وحتى تتم عملية الإسترجاع يقوم المستخدم بكتابة الإستفسار (Query) وسيتم إسترجاع كل المستندات التي لها علاقة بالإستفسار.

وهذه المستندات تحتوي على مجموعة من المواضيع تم جمعها من عدة مصادر وتخزينها بطريقة معينة ويتم ترتيب نتائج البحث المطلوبة علي حسب درجة تناسب الإستفسار مع المستند.

وبناء آ على ذلك يمكن تعريف المدونة بأنها عبارة عن مجموعة من المستندات, هذه المستندات تحتوي على عدة مواضيع تم تخزينها وحفظها بطريقة معينة وتم إستخدام الإستفسارات لتمكن من الوصول إلى البيانات المطلوبة ويجب وجود علاقة بين الكلمات في الإستفسار والكلمات في المستند لكي يتم إسترجاع المستندات المناسبة.

### 1.3 مشكلة البحث :

في السابق كان الباحث الذي يرغب في عمل نظام لإسترجاع المعلومات يجب أولا أن يقوم ببناء مدونة ثم يقوم بإختبار الخوارزمية على هذه المدونة, إذا أراد باحث آخر عمل نظام لإسترجاع المعلومات فإنه أيضا سيحتاج إلى عمل مدونة, إذن المشكلة هي عدم وجود مدونة قياسية للمستخدمين في مجال الإسترجاع, وعدم القدرة على المقارنة بين أنظمة إسترجاع المستخدمين لأن لكل مستخدم مدونة وخوارزمية مختلفة, كما أن العديد من المدونات المتاحة ليست لديها إستفسارات ولا تحتوي على الحكم الملائم, بالإضافة إلى أن تكلفة الحصول على المدونة عالية جدا.

### 1.4 أهمية البحث :

بناء مدونة تكون متاحة لكل المستخدمين في مجال إسترجاع المعلومات. يتمكن المستخدم بواسطة هذه المدونة من إختبار فعالية الخوارزمية وكفاءة نظام الإسترجاع .

### 1.5 أهداف البحث:

يهدف هذا البحث إلى توفير مدونة قياسية باللغة العربية لكل المستخدمين في مجال إسترجاع المعلومات وما يميز هذه المدونة أنها تحتوي على:

1- الحكم الملائم (Relevance Judgment): وهو معرفة المستندات التي يجب أن تسترجع من كل إستفسار.

2- الإستفسارات: تساعد في سرعة الوصول إلى المعلومة المطلوبة .

### 1.6 حدود البحث:

بناء مدونة باللغة العربية متعلقة بمجال الحاسوب .

## 1.7 منهجية البحث:

في هذا البحث تم جمع بيانات متعلقة بمجال الحاسوب ويتم تخزينها في ملفات نصية وحفظها بترميز (UTF-8)

وتم إستخراج كل الإستفسارات التي بواسطتها يتم الوصول إلى البيانات ويتم إستخدام تقنيات مثل ربيدماينر (Rapidmier) لنجري بها العديد من الإحصائيات وبرنامج (Lucene).

## 1.8 هيكلية البحث:

يتضمن البحث الأبواب التالية:

الباب الأول : تعريف المشروع بشكل عام وهدفه ومشاكله.

الباب الثاني : شرح تفاصيل المشروع والدراسات السابقة.

الباب الثالث : الأدوات والتقنيات المستخدمة لتنفيذ المشروع.

الباب الرابع : كيفية بناء وتقييم المدونة.

الباب الخامس : النتائج والتوصيات والمراجع .

## 1.9 خطة البحث:

الجدول (1-1) يوضح خطة البحث

Table الجدول (1-1) يوضح خطة البحث

المهمة	جمع البيانات	الباب الأول	الباب الثاني	الباب الثالث	الباب الرابع	الباب الخامس
الفترة الزمنية	4 أسابيع	أسبوع	3 أسابيع	أسبوعين	4 أسابيع	يومين

# الباب الثاني

الإطار النظري والدراسات

السابقة

## 2 الإطار النظري:

### 2.1 م مقدمة:

ازداد التركيز على أنظمة استرجاع المعلومات خاصة بعد التطورات الحديثة ووصول الإنترنت, والعديد من الدراسات ركزت على إسترجاع المعلومات من المستندات الإنجليزية إلا أن أنظمة إسترجاع المعلومات من المستندات العربية مازالت مجالاتها مهملة, ولكن في السنوات الاخيرة زاد الإهتمام كثيرا بأنظمة إسترجاع المعلومات العربية مما أدى إلى تطوير تقنيات وخوارزميات جديدة [2].

سوف نبدأ البحث بوصف مختصر للغة العربية, وسوف نركز على ميزات اللغة العربية.

### 2.2 خصائص اللغة العربية:

اللغة العربية هي من اللغات العتيقة والقديمة في العالم وهي تنتمي إلى اللغات السامية, بالإضافة إلى أنها لغة الدين الإسلامي ولغة الكتاب المقدس, عدد الحروف العربية 28 حرف, وتعتبر اللغة العربية من اللغات المرنة بمعنى إمكانية إشتقاق عدة كلمات من نفس الجذر [2].

### 2.3 نظام إسترجاع المعلومات:

يعتبر علم إسترجاع المعلومات أحد فروع المعرفة القديمة وهو أقدم من علم الحاسوب وعملية الإسترجاع تمكن المستخدم من الوصول إلى معلومة معينة من مجموعة من البيانات غير المهيكلة, وتتم عملية إسترجاع المعلومات عن طريق إرسال طلب من المستخدم إلى نظام الإسترجاع و تتم معالجة الطلب وإستخلاص وعرض النتائج وفقاً لطلب المستخدم [3].

توجد عدة تعريفات لعملية إسترجاع المعلومات منها تعريف الصوينع بأنه النظام الذي يتعامل مع معلومات وسيطة بين المستفيد والمعلومات النهائية التي يسعى إليها [4].

ويستنتج من التعريفات السابقة أن إسترجاع المعلومات هو البحث عن المعلومات داخل مجموعة كبيرة من المستندات غير المهيكلة وفقاً لطلب المستخدم. هذه البيانات هي عبارة عن نصوص تم

تخزينها بطريقة غير مهيكلة بمعنى أنه لا يوجد شكل أو طريقة معينة يتم بها جمع البيانات وتخزينها, ومن هنا يأتي الفرق بين عملية إسترجاع المعلومات (IR) و قواعد البيانات (DB).

Table الجدول (1-2) يوضح الفرق بين إسترجاع المعلومات و قواعد البيانات

الترقي	المقارنة	IR	DB
1	طريقة جمع البيانات	غير مهيكلة	مهيكلة
2	طريقة تمثيل البيانات	في شكل سلاسل نصية صور, صوت...الخ	في جداول (وأعمدة) شكل صفوف
3	النتائج	غير دقيقة	دقيقة
4	كيفية الوصول إلى البيانات	لا توجد صيغة محددة	توجد صيغة محددة مثلا

Select table table  
; name

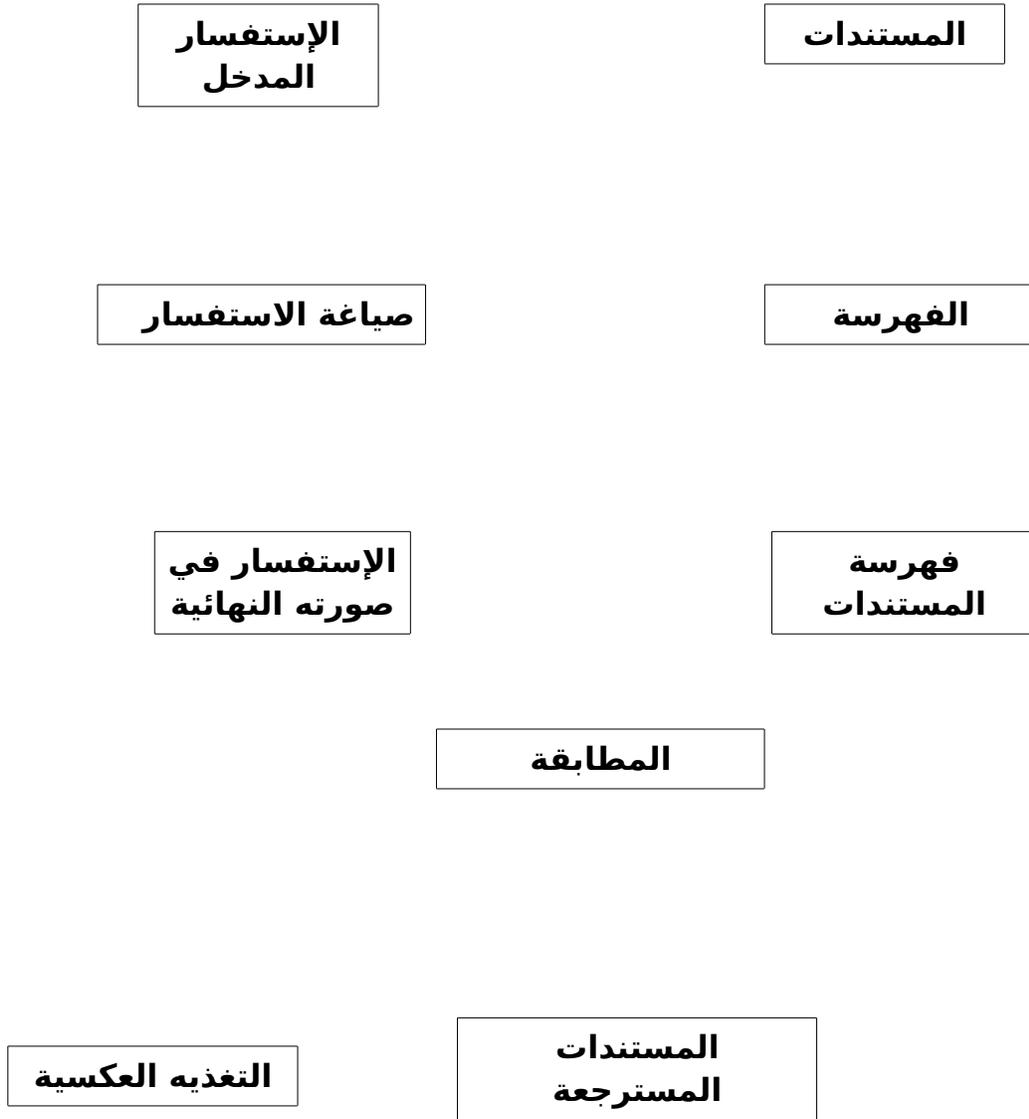
الجدول (1-2) يوضح الفرق بين إسترجاع المعلومات و قواعد البيانات

## 2.4 العمليات الأساسية في إسترجاع المعلومات:

يجب أن تمثل مجموعة المستندات بطريقة معينة تسمى الفهرسة, وتمثل المعلومات المطلوبة في شكل إستفسار لتسهيل عملية الوصول إليها, وعند المقارنة بين الإستفسار والمستند يتم الحصول على النتيجة المطلوبة, وبواسطة عملية المقارنة نحدد توافق المستند مع الإستفسار, ويعتبر المستند مناسب إذا كان وذا لطلب المستخدم [1].

ويقوم نظام الإسترجاع بترتيب المستندات على حسب درجة توافقها مع الإستفسار, بمعنى أن المستند الموجود في الأعلى هو الأكثر تناسبا مع الإستفسار [1].

الشكل (1-2) يشير إلى العمليات الأساسية في إسترجاع المعلومات



الشكل (1-2) يوضح العمليات الأساسية في إسترجاع المعلومات

## 2.5 كيفية تقييم نظام إسترجاع المعلومات:

إن تقييم نظام الإسترجاع ضروري لقياس نجاح النظام ومدى تحقيقه للأهداف, ويوجد معيارين لقياس كفاءة النتائج وهما معدل الدقة ومعدل الإستدعاء.

### 2.5.1 معدل الدقة (Precision Rate):

عدد المستندات ذات الصلة المسترجعة على عدد المستندات الكلي المسترجع.

A	C
---	---

C: number of irrelevant documents retrieved (عدد المستندات الغير مناسبة المسترجعة)

A: number of relevant documents retrieved (عدد المستندات المناسبة المسترجعة)

ويمثل معدل الدقة كالآتي:

$$\frac{A}{A+C} * 100\%$$

عدد المستندات المناسبة المسترجعة \* 100%  
عدد المستندات المناسبة المسترجعة + عدد المستندات الغير مناسبة المسترجعة

## 2.5.2 معدل الإستدعاء (Rate Recall)

عدد المستندات ذات الصلة المسترجعة على عدد المستندات ذات الصلة الكلي.

B: number of relevant documents not retrieved ( عدد المستندات المناسبة الغير المسترجعة )

A: number of relevant documents retrieved ( عدد المستندات المناسبة المسترجعة )  
ويمثل معدل الاستدعاء كالآتي:

$$\frac{A}{A+B} * 100\%$$

$$\%100 * \frac{\text{عدد المستندات المناسبة المسترجعة}}{\text{العدد الكلي للمستندات المسترجعة}} =$$

تتمناؤنؤء

B A

المثال يوضح مفهومي الدقة والاستدعاء:

لنفرض أن لدينا مدونة تحتوي 100 مستند، وكان عدد المستندات المناسبة التي يسترجعها الإستفسار

الأول 14 مستند، وبعد إجراء عملية الإسترجاع تم إسترجاع 10 مستند منها 7 مستندات مناسبة و 3 مستندات غير مناسبة، أحسب معدل الدقة ومعدل الإستدعاء؟

الحل

عدد المستندات المفترض أن يسترجعها الإستفسار = 14 مستند

عدد المستندات المسترجعة = 10 مستند

عدد المستندات المناسبة = 7 مستند

عدد المستندات الغير مناسبة = 3 مستند

معدل الدقة =

\*

عدد المستندات المناسبة المسترجعة

%100

عدد المستندات المناسبة المسترجعة + عدد المستندات الغير مناسبة المسترجعة

$$7 \times \%100 = \%70$$

10

معدل الإستدعاء =

%100 \*

عدد المستندات المناسبة المسترجعة

العدد الكلي للمستندات المسترجعة

$$7 \times \%100 = \%50$$

14

## 2.6 نماذج إسترجاع المعلومات:

توجد عدة نماذج لإسترجاع المعلومات التي تختلف فيما بينها بطريقة تمثيل الوثائق والإستفسارات والمطابقة والترتيب ويمكن تصنيف هذه النماذج إلى:

## 2.6.1 نموذج البحث البولياني (Boolean Search Model):

يعرف عادة بنموذج التطابق التام (Exact Match), وفي هذا النموذج يتم صياغة الإستفسار في شكل مجموعة كلمات (عبارات) مع عمليات منطقية مثل ( او OR , و AND , نفي NOT ) ويعتمد هذا النموذج على مقارنة مصطلحات الإستفسار البولياني مع المصطلحات المستخدمة لتمثيل محتوى المستندات، ليحدد ما إذا كان مصطلح الإستفسار موجود أو غير موجود ضمن المصطلحات التي تمثل المستند، ولا بد أن تكون المصطلحات الخاصة بمجموعة المستندات مطابقة لتلك التي ترد في إستفسار البحث حتى يتم إسترجاعها ومن عيوب هذا النموذج أنه لا يسمح بترتيب بيانات المستندات المسترجعة حسب أهميتها و صلتها بالموضوع، فالمواد المسترجعة تعد على نفس الدرجة من الأهمية. [1]

بالإضافة إلى أن النموذج البولياني ليس طريقة مفضلة للمستندات المتتالية أو المتتابعة بمعنى آخر أنه لا يدعم ترتيب المستندات لهذه الأسباب تعتبر الأقل شهرة من النماذج الأخرى، يتطلب أن يكون المستخدم لديه معرفة عن العمليات المنطقية، بناء على ذلك ليس كل المستخدمين قادرين على الحصول على النتائج المرضية مع هذا النموذج، وأيضاً يستهلك زمن لإسترجاع كل المستندات الملائمة في حالة البيانات الكبيرة .

## 2.6.2 نماذج الترتيب (Ranked Retrieval Models):

## نموذج معالجة حيز المتجهات

### (The Vector Space Processing Model)

يعمل على تحديد درجة التشابه بين المستندات والإستفسارات من خلال قياس المتجهات للمصطلحات ويمكن وفقاً لهذا النموذج الحكم على تشابه المستندات بحساب المصطلحات المتشابهة فيها، كما يمكن إحتساب حيز المتجهات في الإستفساروم قارنته بحيز المتجهات في المستندات المخزنة لتحديد التشابه بينهم [1].

$$\frac{\vec{d}_j \cdot \vec{q}}{|\vec{d}_j| \cdot |\vec{q}|} = \frac{\sum_{i=1}^t (w_{ij} \cdot w_{iq})}{\sqrt{\sum_{i=1}^t w_{ij}^2 \cdot \sum_{i=1}^t w_{iq}^2}}$$

:Dj المستند في المتجه

:Q الاستفسارات في المتجه

:Wij وزن العنصر في المستند

:Wiq وزن العنصر في الاستفسار

## نموذج الإسترجاع الإحتمالي

### Model : (Probabilistic Retrieval )

هذا النموذج أساسه إستخدام نظرية الإحتمالات, فبدلاً من مطابقة نفس المصطلحات الواردة في إستفسار البحث مع المصطلحات الواردة في وصف المستندات, فإنه يتم وفقاً للنموذج الإحتمالي إحصاء الإحتمالات التي يمكن أن تكون فيها المستندات ذات صلة بمستخدم معين, ومن ثم ترتب المستندات المسترجعة ترتيباً تنازلياً وفقاً لإحتمالات صلتها بالإستفسار وفائدتها بالنسبة للمستخدم [5] .

$$p(R | x) = \frac{p(x | R)p(R)}{p(x)}$$

$$p(NR | x) = \frac{p(x | NR)p(NR)}{p(x)}$$

X: مجموعة المستندات

R: مجموعة المستندات المسترجعة

NR: مجموعة المستندات الغير مسترجعة

## 2.7 معالجة النصوص في لغات إسترجاع المعلومات:

### **Text Processing In Language Information)** **((Retrieval ( CLIR**

#### 2.7.1 التجزئة (Tokenization) :

هي عملية إستخراج العبارات والكلمات من المستندات وحذف علامات الترقيم والرموز الخاصة بالكلمة العربية أو الإنجليزية.

#### 2.7.2 الكلمات المهملة (Stopword)

هي الكلمات ذات التكرار العالي مثل حروف الجر ويتم إهمالها في أغلب محركات البحث، ولأن هذه الكلمات تتكرر كثيرا فإن إهمالها يؤدي الى تقليل حجم فهرس البحث .

#### 2.7.3 التطبيع (Normalization):

يهدف إلي حذف علامات الترقيم و يحول النص إلى عدة أشكال, توجد القليل من التقنيات تدعم التطبيع باللغة العربية.

#### 2.7.4 التجريد ( Stemming) :

هي تقنية شائعة لكثير من محركات البحث وهي عملية تتم بواسطتها تمثيل مختلف الكلمات التي لها نفس الجذر ويساعد في تقليل حجم الفهرسة.

## 2.8 الدراسات السابقة السابعة:

### 2.8.1 الدراسة الأولى:

مدونة الخليج (مراد عباس, 2004) تتكون من 1000 مقالة تم جمعها من الجرائد عبر شبكة الإنترنت, وهذه المقالات تحتوي على أكثر من 5000 مستند أي قد يصل عدد كلماتها إلى 3 مليون كلمة تقريباً, وتم التخلص من كلمات المهملة, وهي تحتوي على 4 مواضيع (أخبار عالمية عدد مستنداتها 953 مستند, وأخبار محلية عدد مستنداتها 2398 مستند, و مواضيع إقتصادية عدد مستنداتها 909 مستند, ومواضيع رياضية عدد مستنداتها 1430 مستند) أي أن مجموع المستندات هي 5690 مستند, ومن عيوبها لا تحتوي على الحكم الملائم والاستفسارات.

### 2.8.2 الدراسة الثانية:

مدونة الوطن (مراد عباس, 2004) تتكون من 20000 مقالة وتحتوي على 6 مواضيع وهي مواضيع بيئية عدد مستنداتها 2782 مستند, ومواضيع دينية عدد مستنداتها 3860 مستند, ومواضيع إقتصادية عدد مستنداتها 3468 مستند, وأخبار محلية عدد مستنداتها 3596 مستند, وأخبار عالمية عدد مستنداتها 2035 مستند, ومواضيع رياضية عدد مستنداتها 4550 مستند, اي ان عدد المستندات هي 20291 مستند, ومن عيوبها لا تحتوي على الحكم الملائم والاستفسارات.

### 2.8.3 الدراسة الثالثة:

مدونة قيس ديوكس (Quranic Arabic Corpus, 2009) يهتم بقواعد اللغة العربية مثل النحو والصرف ويتم التحليل على ثلاثة مستويات وهي الصرف والنحو والأنطولوجيا, يتكون من 77430 كلمة يهدف إلى توضيح قواعد الصرف والنحو للباحثين في دراسة لغة القرآن الكريم.

# الباب الثالث

الأدوات والتقنيات

### 3 الأدوات والتقنيات المستخدمة:

#### 3.1 ريد ماينر (Rapidminer)

يوجد هذا البرنامج بشكل مستقل لتحليل البيانات وتعطي مستخدميه ميزة تنافسية عالية، و من مميزات البرنامج تكامل البيانات و تحليلها و إعداد التقارير و تمتعه بواجهة رسومية سهلة الإستخدام لتصميم عمليات التحليل، ويحتوي على مستودعات للعمليات و البيانات و معالجة البيانات الوصفية، يدعم التعرف على الأخطاء بسرعة ويعالجها بسرعة، و يعتبر برنامج كامل و مرن يمكنه تحميل مجموعه من البيانات و تحويلها و نمذجتها و تصويرها، بالإضافة الى أنه برنامج مجاني مفتوح المصدر [7].

خطوات عمل برنامج الرييد ماينر:

## 1- تحديد مسار المستندات :

يتم فيه توضيح المسار التي توجد فيه المستندات.

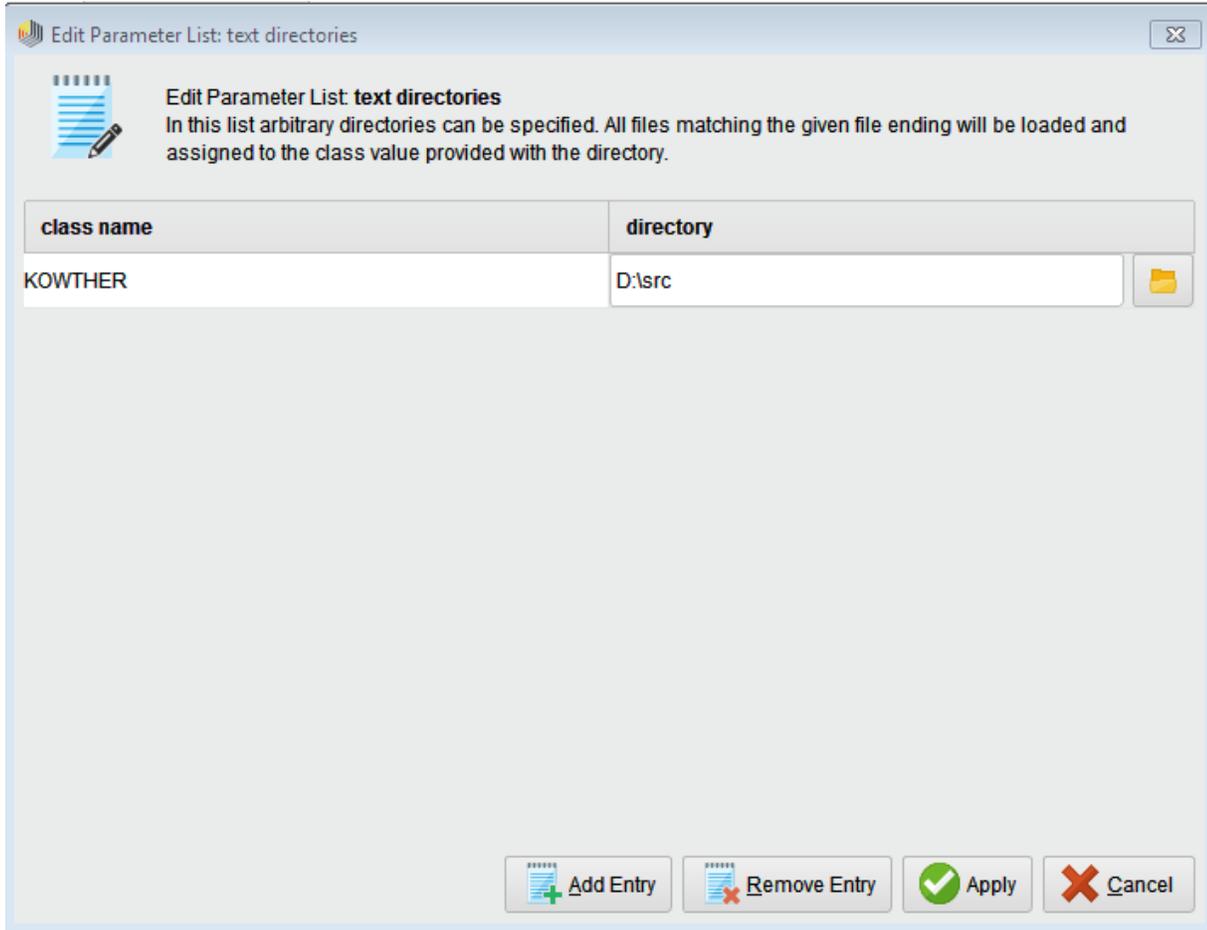


Figure الصورة يوضح (1\_3) تحديد المسار لتنفيذ برنامج rapidminer

2- معالجة البيانات الموجودة في الملفات (process document from file):

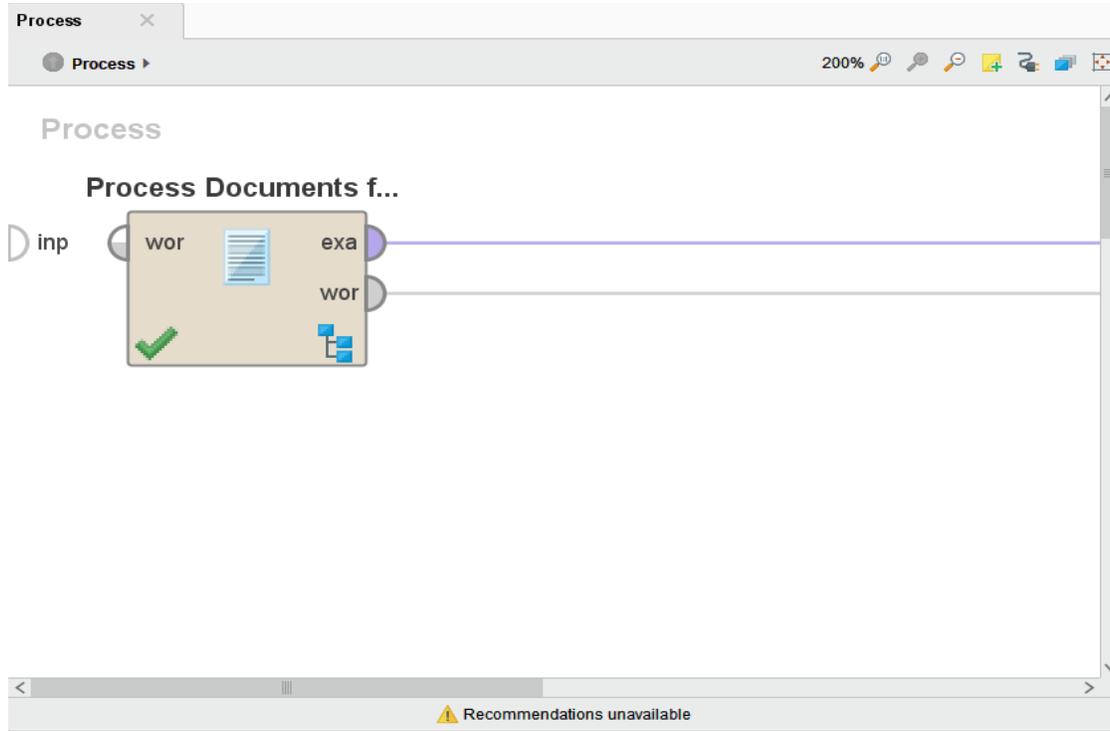


Figure الصورة (2\_3) يوضح معالجة البيانات

### 3- عمل تجزئه (Tokenize) للمستندات :

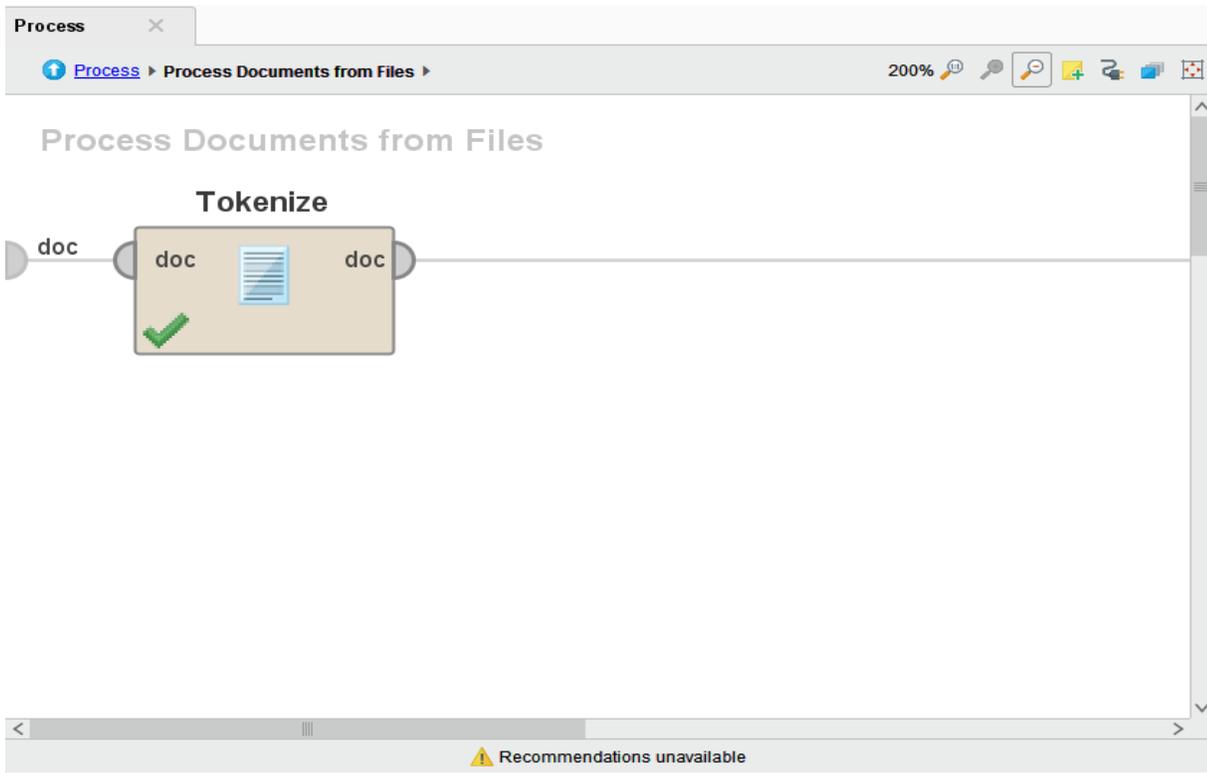


Figure الصورة (3\_3) يوضح عمل Tokenize للمستندات

4- عند تنفيذ البرنامج يظهر قائمة تحتوي الكلمات الموجودة في المستندات و قائمة الكلمات الفريدة :

Word	Attribute Name	Total Occurrences	Document Occurrences	jk
لئلي	لئلي	1	1	1
لئلية	لئلية	2	2	2
لئلين	لئلين	1	1	1
لئلوب	لئلوب	1	1	1
لئلوب	لئلوب	1	1	1
لئفاس	لئفاس	1	1	1
لئرنا	لئرنا	1	1	1
لئصح	لئصح	4	3	4
لئصحت	لئصحت	7	3	7
لئصلا	لئصلا	2	2	2
لئضافة	لئضافة	1	1	1
لئضحي	لئضحي	1	1	1
لئعظم	لئعظم	1	1	1
لئعمال	لئعمال	3	1	3
لئعمالها	لئعمالها	1	1	1

Figure الصورة (4\_3) يوضح تنفيذ البرنامج

5- يستخدم برنامج الإكسل (Microsoft Excel) لحساب عدد الكلمات الموجودة في المستندات :

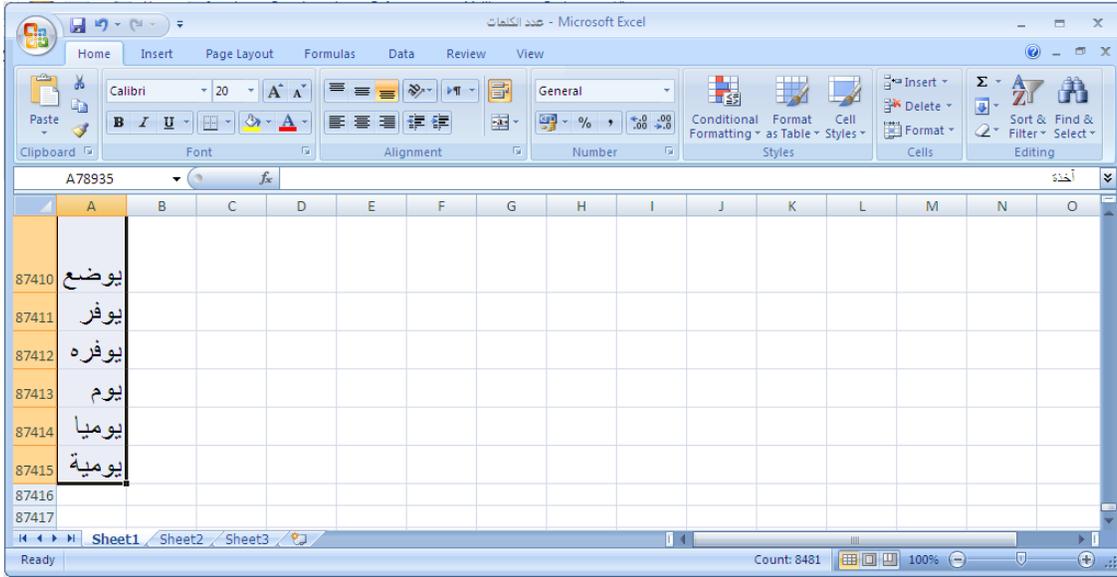


Figure الصورة (5\_3) يوضح حساب عدد الكلمات في المستندات

6- ويتم ايضا حساب عدد الكلمات الفريدة (Total Occurences) الموجودة في المستندات:

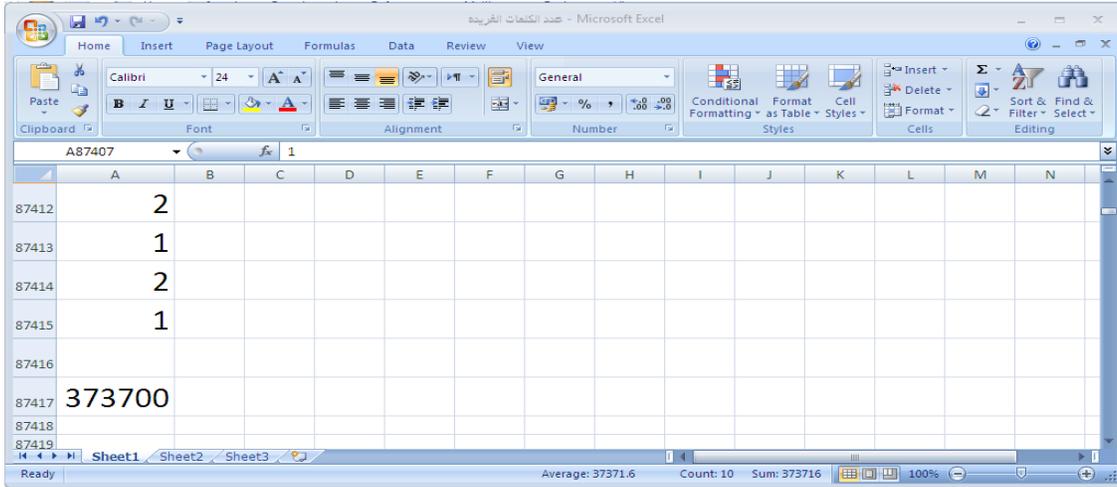


Figure الصورة (6\_3) يوضح حساب عدد الكلمات الفريدة في المستندات

## 3.2 الجافا (Java Development Kit (JDK)

تعد لغة الجافا لغة سهلة لإعداد التطبيقات المختلفة، بسبب احتوائها على العديد من الميزات التي تسهل على المبرمج عمله حيث يمكنه من تطبيق برمجيات خالية من الأخطاء، تعمل الجافا على عدة أنظمة تشغيل كالويندوز واليونكس وغيرها، ويتم معالجة التطبيقات والبرامج في الجافا بصورة سريعة. [12]

## 3.3 ++Notepad

هو محرر نصوص لكتابة الاكواد يدعم مختلف البرامج ولغات البرمجة يمكنه فتح أكثر من ملف نصي في وقت واحد، كذلك تسمح للمستخدمين بتحميل عملهم الى مواقع الأنترنت، يحتوي على محرك بحث متقدم يتميز بصغر حجمه وتوافقه مع أغلب إصدارات الويندوز، ويمكن الحصول عليه من المصادر المفتوحة. [13]

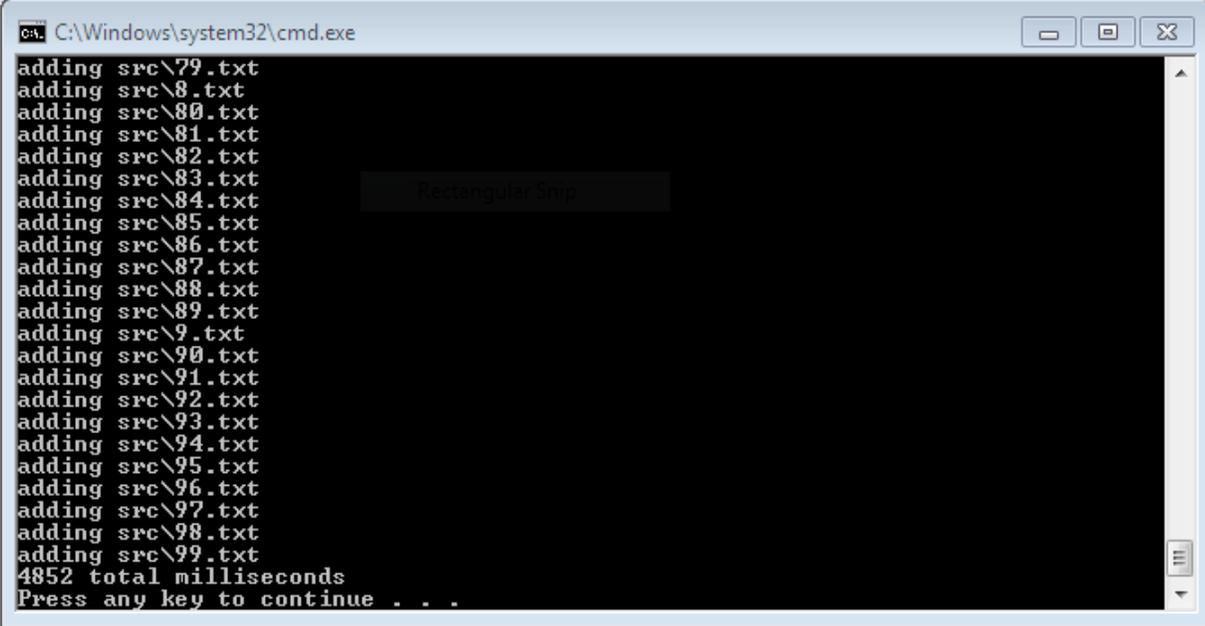
## 3.4 برنامج (Lucene):

اللوسين هي عبارته عن مكتبة لعمليتي الفهرسة والبحث في الملفات النصية، وبرنامج اللوسين مكتوب بلغة الجافا المتاحة والمفتوحة المصدر وهو ليس تطبيق قائم بذاته وهو مصمم بطريقة مرنة ليتعامل مع التطبيقات بسهولة [9]

الميزات:  
قابل للتطوير وذات أداء عالي، الدقة وكفاءة البحث، يعتبر أساس لبرمجيات الأباتشي ( Apache )  
[[10]

خطوات العمل:

1- في عملية الفهرسة يتم عرض كل المستندات .



```
C:\Windows\system32\cmd.exe
adding src\79.txt
adding src\8.txt
adding src\80.txt
adding src\81.txt
adding src\82.txt
adding src\83.txt
adding src\84.txt
adding src\85.txt
adding src\86.txt
adding src\87.txt
adding src\88.txt
adding src\89.txt
adding src\9.txt
adding src\90.txt
adding src\91.txt
adding src\92.txt
adding src\93.txt
adding src\94.txt
adding src\95.txt
adding src\96.txt
adding src\97.txt
adding src\98.txt
adding src\99.txt
4852 total milliseconds
Press any key to continue . . .
```

Figure الصورة (7\_3) يوضح كيفية فهرسه المستندات في برنامج (Lucene)

2- في عملية البحث يتم إدخال الإستفسار وبعد ذلك يتم تجريد الإستفسار من الحروف الزائدة :

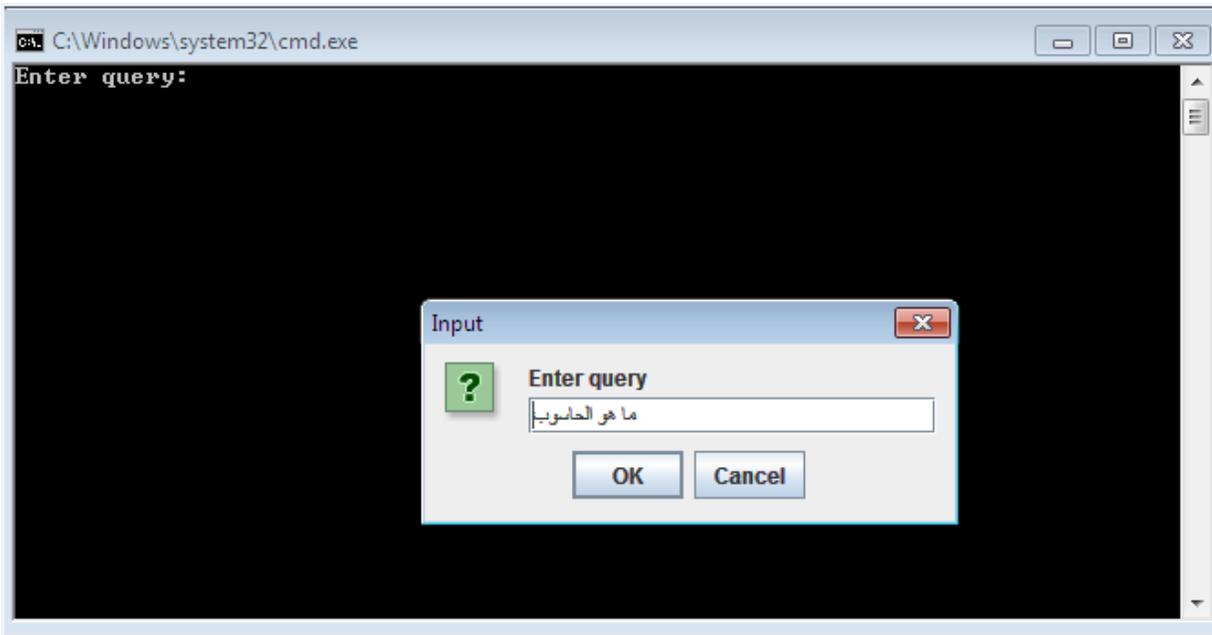


Figure الصورة (8\_3) يوضح عملية إدخال الإستفسار في برنامج (Lucene)

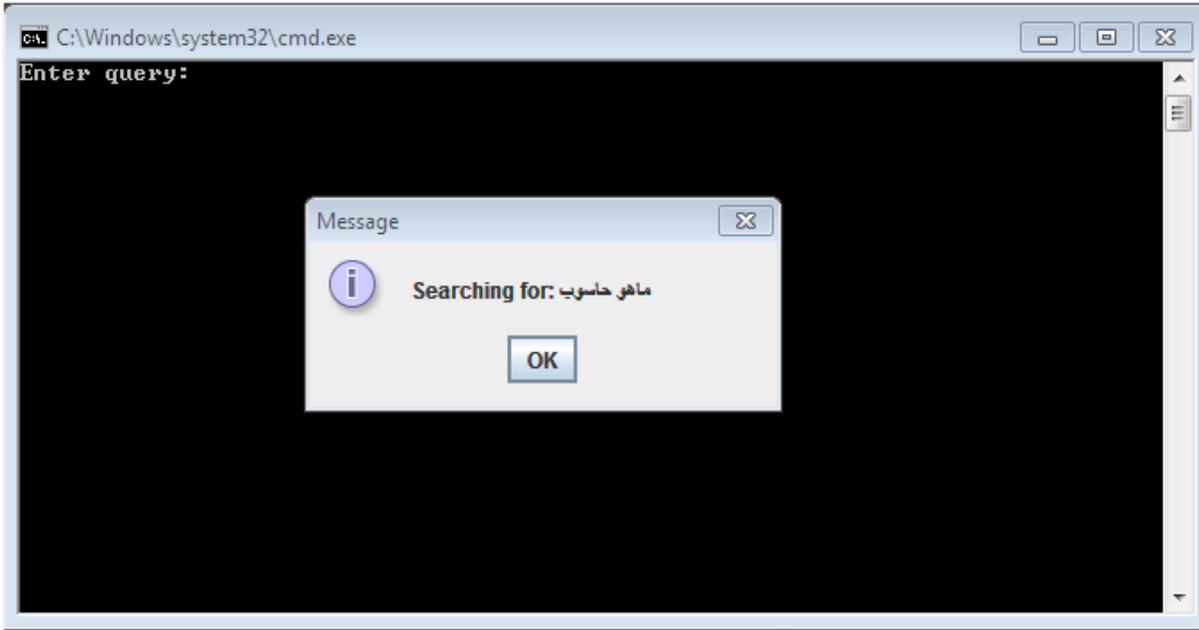


Figure (Lucene) الصورة(9\_3) يوضح عملية البحث في برنامج

3- وعند التنفيذ بالضغط **OK** على يعرض المستندات ذات الصلة بالإستفسار

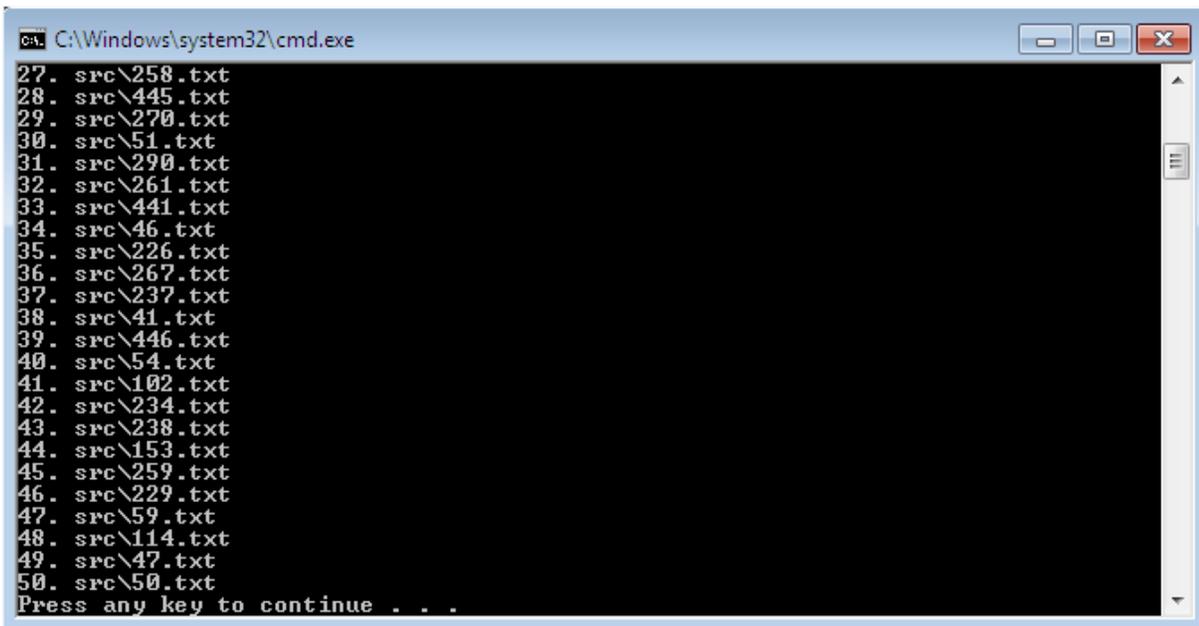


Figure الصورة(10\_3) يوضح عملية التنفيذ إسترجاع الملفات

# الباب الرابع

بناء المدونة

## 4 الباب الرابع

### 4.1 المقدمة:

توجد العديد من الدراسات التي تناولت مفهوم استرجاع المعلومات, ولكن كل دراسته كانت تأخذ الموضوع من وجهه نظر مختلفه عن الدرسته الاخرى لان مفهوم استرجاع المعلومات مجال واسع ومتطور, وكل دراسته لها هدف وطريقه معينه و قد تتداخل الدراسات في بعض المفاهيم والمشاكل وغيرها .

### 4.2 المدونة (Corpora) :

من الصعوبات التي يجدها الباحث في مجال استرجاع المعلومات باللغه العربيه هو عدم توفر مدونه يستطيع بواسطتها استرجاع المعلومات, ويجب ان تحتوي كل مدونه على الاتي:

1- مجموعه من المستندات (Documents).

2- مجموعه من الاستفسارات (Queries).

3- الحكم الملائم (Relevance Judgment) :

وهو معرفه المستندات التي يجب ان تسترجع من كل استفسار, و يقوم باسترجاع جميع المستندات ذات الصله بالاستفسار مع عدم استرجاع المستندات التي ليست لها صله.

### 4.3 مجموعة المستندات:

تتكون هذه المستندات من مواضيع مختلفه تم جمعها من عده مصادر مثل الكتب والمقالات وشبكه الانترنت, بعد جمع هذه البيانات تم حفظها في ملفات نصيه وتم اجراء بعض المعالجات

مثل حذف الكلمات الانجليزيه والتخلص من الرسومات والجداول, وتم حفظ البيانات بترميز (UTF-8).

ان اختيار هذه المواضيع كان على اساس معين وهي ان تكون المواضيع متعلقه بمجال الحاسوب  
لانه يعتبر المجال الاقرب والاساسي للباحث فمثلا اذا اعطي الباحث استفسار فانه يستطيع تحديد  
المستندات المناسبه والغير مناسبه للاستفسار بحكم معرفته وخبرته بمجال الحاسوب ولكن ان اعطي  
استفسار يخص مجال آخر (طبي,هندسه,...الخ) فان معرفته تكون اقل ولايستطيع تحديد المستندات.

توجد عدة ادوات استخدمها الباحث مثل برنامج الراييد ماينر (Rapidminer) والاكسل (Excel)  
وهذه الادوات ساعدت الباحث على الحصول مجموعته من الاحصائيات المهمه

والجدول (4-1) يوضح الاحصائيات:

Table الجدول (1-4) يوضح احصائيات الرييد ماينر (Rapidminer)

المدونة	عددتها
عدد الملفات النصية	600
عدد المواضيع	40
الكلمات	87415
الإستفسارات	313
متوسط طول المستند	307
الكلمات الفريدة	373696

الجدول (4-1) يوضح احصائيات الرييد ماينر (Rapidminer)

## 4.4 مجموعة الإستفسار:

هي عبارة عن مجموعة من الاسئلة عن البيانات الموجودة في المستندات, ويجب ان تكون هناك  
علاقة بين الكلمات في الاستفسار والكلمات في المستند حتى تتم استرجاع المستندات المناسبه, لان  
الهدف من الفهرسة هو تسريع عمليه البحث عن المعلومة التي يطلبها المستخدم خاصة عندما تكون  
البيانات كبيره.

أن عدد المستندات في هذا البحث هو 40 مستند وهذا العدد يعتبر كافي لعمل نظام استرجاع معلومات وتقييمه والوثوق من نتائجه, وذلك بناءً على دراسته في علم الاحصاء وهي اذا اراد باحث عمل تجربته فانه لا ينفذه على المجتمع كله إنما يأخذ عينة منه , واصغر عينة يمكن استخدامها هي 32 عينة, واذا كانت العينة اقل من 32 فانها لا تعني شيئاً ونتائجها غير مهمة وغير موثوقة, واذا كانت العينة أكثر من 32 فانها تعتبر كافية وذات اهمية [14].

#### 4.4.1 كيف تم بناء المدونة:

تم بناء هذه المدونة عن طريق إختيار مجموعة من المواضيع المتعلقة بمجال الحاسوب, قد نطلق على المدونة مصطلحات اخرى مثل الذخائر النصية أو مجموعة نصوص وهي عبارة عن مجموعة من المستندات , تلك المستندات تحتوي على مجموعة من المواضيع تم الحصول عليها من عدة مصادر مثل (شبكة الأنترنت, كتب, مراجع, مقالات) وبعد ذلك تم وضع النصوص التي تخص موضوع واحد في ملفات نصية (Text Document) وتم حفظه بترميز (UTF-8), وتم إجراء عدة معالجات للمواضيع قبل حفظها ومنها

تم حذف الكلمات الإنجليزية التي لا تؤثر على معنى الجملة ولكن التي لها تأثير على الجملة تم ترجمتها بأفضل طريقة الى اللغة العربية وتم تحويل الإختصارات الإنجليزية الى حروف عربية مثلًا (CRM) الى (سي آر أم), وايضا تم حذف الجداول والصور .

تم فتح كل ملف نصي وإستخراج عدد من الإستفسارات تلك الإستفسارات تمكن المستخدم من الوصول الى المستند, والإستفسار هي عبارة عن أسئلة, ومعرفة المستندات التي يجب أن تسترجع من كل إستفسار ذلك ما يسمى بالحكم الملائم (Relevance Judgment).

ومن الصعوبات التي واجهت الباحث التكلفة والجهد في جمع البيانات, تم استخدام عدة برامج ومنها برنامج (Rapidminer) حيث يقوم الباحث بإدخال جميع البيانات على البرنامج لتتم معالجتها, وعند إدخال البيانات يتم فحصها وتجزئتها (Tokenize) ومن ثم حساب عدد الكلمات الموجودة في المستندات والكلمات الفريدة (Total Occurences) بواسطة برنامج الإكسل (Excel) .

وبعد اكتمال تلك العمليات يقوم الباحث بإدخال الإستفسارات في برنامج (lucene) الذي يعمل على إزالة الحروف الزائدة مثل (ال) وغيرها , ففي البداية يتم عملية الفهرسة لجميع المستندات الموجودة وبعد ذلك يتم البحث عن المعلومات من أجل إسترجاع المستندات التي لها علاقة بالإستفسار .



# الباب الخامس

النتائج والتوصيات

## 5 النتائج والتوصيات:

### 5.1 النتائج:

1- توفير مدونة لكل المستخدمين في مجال الحاسوب.

### 5.2 التوصيات:

- 1- توسيع مدونة باضافة مستندات ومجموعة من الإستفسار ومعرفة الحكم الملائم (Relevance Judgment).
- 2- بناء مدونة لإستخدامها في اللهجات العربية.
- 3- بناء مدونة تشمل المجالات العامة.
- 4- حفظ المدونة في صيغة (XML) لإن فيه يحدد العنوان (Title) والمحتوى (Body) فذلك يساعد المستخدمين في الإسترجاع بالعناوين دون الدخول في المحتوى.
- 5- إن علم إسترجاع المعلومات من العلوم الخصبة والمتطورة فيجب البحث عن أدوات وتقنيات جديدة وانشاء خوارزميات وذلك بهدف تلبية احتياجات المستخدم.

### 5.3 الخاتمة:

قد تتأثر الطريقة المستخدمة لاجراء البحث في النتائج النهائية ولا يوجد مشروع كامل تكون نتائجها مؤكدة ومضمونة فلكل مشروع نواقص وهذا المشروع ليس استثناء .

المراجع

[1] Ali Mohammed Mustafa, Mixed-Language Arabic- English Information Retrieval,

(February 2013, University of Cape Town.

[2] Zainab Majeed , "Search Queries in an Information Retrieval System for Arabic-Language Texts" , (2014), University of Kentucky.

[3] فاتن سعيد, (يوليو 2007), استرجاع المعلومات في المكتبات الرقمية , الصفحات(1).

[4] نادية مصطفى العيدورس أحمد, (يونيو 2007), استخدام اللغة العربية في نظم استرجاع المعلومات, الصفحات(166).

[5] د. فاتن سعيد بامفلح , (2006), نماذج استرجاع المعلومات الصفحات (41 - 42 - 52 - 53 - 55) .

[6]

[file:///D:/New%20folder/END/project%20final/step%20number%201reading/  
%D8%A7%D9%84%D8%B1%D9%82%D9%85%D9%91%D9%8A  
%D8%A7%D8%AA%20\\_  
%20%D9%85%D9%82%D8%A7%D9%84%D8%A7%D8%AA%20\\_  
%20%D8%B0%D9%83%D8%A7%D8%A1%20%D8%B5%D9%86%D8  
%B9%D9%8A%20\\_%20%D9%84%D9%85%D8%AD  
%D8%A9%20%D8%B9%D9%86%20%D9%86%D8%B8%D9%85%20  
%D8%A7%D8%B3%D8%AA%D8%B1%D8%AC  
%D8%A7%D8%B9%20%D8%A7%D9%84%D9%85%D8%B9%D9%84  
%D9%88%D9%85%D8%A7%D8%AA  
%20%D8%A7%D9%84%D8%B6%D8%A8%D8%A7%D8%A8%D9%8A  
%D8%A9%20Fuzzy%20Information%20Retrieval.html](file:///D:/New%20folder/END/project%20final/step%20number%201reading/%D8%A7%D9%84%D8%B1%D9%82%D9%85%D9%91%D9%8A%D8%A7%D8%AA%20_%20%D9%85%D9%82%D8%A7%D9%84%D8%A7%D8%AA%20_%20%D8%B0%D9%83%D8%A7%D8%A1%20%D8%B5%D9%86%D8%B9%D9%8A%20_%20%D9%84%D9%85%D8%AD%D8%A9%20%D8%B9%D9%86%20%D9%86%D8%B8%D9%85%20%D8%A7%D8%B3%D8%AA%D8%B1%D8%AC%D8%A7%D8%B9%20%D8%A7%D9%84%D9%85%D8%B9%D9%84%D9%88%D9%85%D8%A7%D8%AA%20%D8%A7%D9%84%D8%B6%D8%A8%D8%A7%D8%A8%D9%8A%D8%A9%20Fuzzy%20Information%20Retrieval.html)

[7]

<http://www.freshestsoft.com/2012/09/rapidminer-52008.html>

التاريخ 04 /10/ 2016 الزمن AM 11:06

[8]

<http://std.ena.gov.lb/Files/Samples/DataMiningAndForecasting.pdf>

التاريخ 10 /2016/4 الزمن AM 11:19

[9]

<http://www-pagines.fib.upc.es/~ri/session1-eng.html>

التاريخ 2016/10/04 الزمن 12:05AM

[10]

<http://lucene.apache.org/core>

التاريخ 2016/10/04 الزمن AM 12:31

[11]

<http://vb4arb.com/vb/thread-5495.html>

التاريخ 2016/10/04 الزمن PM 1:13

[12]

[http://mawdoo3.com/%D9%85%D8%A7\\_%D9%87%D9%8A\\_%D9%84%D8%BA%D8%A9\\_%D8%A7%D9%84%D8%AC\\_%D8%A7%D9%81%D8%A7](http://mawdoo3.com/%D9%85%D8%A7_%D9%87%D9%8A_%D9%84%D8%BA%D8%A9_%D8%A7%D9%84%D8%AC_%D8%A7%D9%81%D8%A7)

[13]

<http://soft.sptechs.com/74565-Notepad++.html>

[14]

file:///C:/Users/TOSHIBA/Desktop/New%20folder%20(6)/%D9%82%D9%88%D8%A7%D8%B9%D8%AF%20%D9%85%D9%8A%D8%B3%D8%B1%D8%A9%20%D9%81%D9%8A%20%D8%A7%D8%AE%D8%AA%D9%8A%D8%A7%D8%B1%20%D8%AD%D8%AC%D9%85%20%D8%A7%D9%84%D8%B9%D9%8A%D9%86%D8%A9.html



الملاحق

الجدول توضح إستخراج الإستفسارات ومعرفة الحكم الملائم (Relevance Judgment):

## 1-نظام المعلومات

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هو نظام المعلومات؟

اذكر امثله على نظم المعلومات؟

كيف نشأ نظم المعلومات؟

ما هي تطبيقات نظم المعلومات؟

ما هي مجالات العمل في نظم

المعلومات؟

ماهو المقصود بنظم المعلومات

الحاسوبية؟

ما هو الاختلاف بين نظم المعلومات

وتخصص علوم الحاسوب؟

اهميه نظم المعلومات؟

انواع نظم المعلومات؟

ماهي علاقه البيانات بالمعلومات؟

كيف اصبحت نظم المعلومات تمثل

جانبا من حياتنا؟

لماذا نظم المعلومات؟

ما هي تحديات نظم المعلومات؟

10.txt 9.txt 8.txt 7.txt 6.txt 5.txt 4.txt 3.txt 2.txt 1.txt

15.txt 14.txt 13.txt 12.txt 11.txt

## 2-النظم الخبيره

الاستفسار

ما هي النظم الخبيره؟

ماهي خصائص ومميزات النظم الخبيره؟

ماهي مكونات النظم الخبيره؟

الملفات النصيه التي يجب ان تسترجع

22.txt 21.txt 20.txt 19.txt 18.txt 17.txt 16.txt

29.txt 28.txt 27.txt 26.txt 25.txt 24.txt 23.txt

اشهر النظم الخبيره؟

ماهي الفرق بين قواعد البيانات و النظم الخبيره؟

اذكر امثله على النظم الخبيره؟

ماهي مراحل النظم الخبيره؟

30.txt

3-الحوسبه الموزعه

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ماهي الحوسبه الموزعه؟

ماهي مميزات الحوسبه الموزعه؟

ماهي الانظمه الموزعه؟

37.txt 36.txt 35.txt 34.txt 33.txt 32.txt 31.txt

ماهو الفرق بين انظمه الكمبيوتر الموزعه

والمتزامنه؟

ما هي انظمه التشغيل الموزعه؟

44.txt 43.txt 42.txt 41.txt 40.txt 39.txt 38.txt

ماهي الحوسبه العنقوديه؟

ممايتألف الانظمه الموزعه؟

متى بدأت الانظمه الموزعه؟

ماهي التطورات في الانظمه الموزعه؟

45.txt

كيف يتم التعاون بين العمليات في الانظمه

الموزعه؟

#### 4-تفاعل الانسان والحاسوب

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هو تفاعل الانسان الحاسوبي؟

ماهي العوامل الرئيسيه التي يجب ان يؤخذ في

الاعتبار عند تصميم التفاعل؟

52.txt 51.txt 50.txt 49.txt 48.txt 47.txt 46.txt

ماهو مستقبل تفاعل الانسان والحاسوب؟

كيف يفكر الانسان وكيف يعمل الحاسوب؟

ما هي اهداف تفاعل الانسان والحاسوب؟

59.txt 58.txt 57.txt 56.txt 55.txt 54.txt 53.txt

ما هو المقصود بسهولة الاستخدام؟

ماهو مبادئ تصميم واجهه المستخدم؟

لماذا تفاعل الانسان والحاسوب مهم؟

60.txt

#### 5-قضايا اخلاقيه ومهنيه

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هي المبادئ التي يلتزم بها مهندسوا

البرمجيات لرفاهيه المجتمع؟

67.txt 66.txt 65.txt 64.txt 63.txt 62.txt 61.txt

ما هي وثيقه قواعد السلوك والاخلاق المهنيه؟

ما هي اقسام القيم والمبادئ المهنيه؟

74.txt 73.txt 72.txt 71.txt 70.txt 69.txt 68.txt

ماهي سلوك الزملاء الاخلاقيه؟

ماهو مفهوم التوجيه الارشادي؟

كيف يتم توجيه المنظومه الاخلاقيه؟

ما هي نصوص الميثاق؟

ما هي السلوك الاخلاقي للاعلاميين؟

75.txt

## 6-طرق بحث

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هو البحث العلمي؟

متى نشأ البحث العلمي؟

خطوات اعداد البحث العلمي؟

ما هي اساليب البحث العلمي؟

ما هو البحث؟

82.txt 81.txt 80.txt 79.txt 78.txt 77.txt 76.txt

ما هي مرحله تصميم بحث الطرق المختلطه؟

كيف يدخل البحث العلمي في كافه انواع الحقول

المعرفيه؟

89.txt 88.txt 87.txt 86.txt 85.txt 84.txt 83.txt

ماهي مستويات البحوث؟

اذكر مهارات ضروريه في كتابه البحث العلمي؟

ما هي صفات الباحث؟

ماهو المنهج العلمي ؟

90.txt

ماهي المصطلحات المفتاحيه من تعريف البحث؟

## 7-المسجلات

### الاستفسار

ما هي المسجلات؟

فيما تستخدم المسجلات؟

ما انواع المسجلات؟

ماهي مسجلات الاغراض الخاصه؟

ماهي مسجلات الاغراض العامه؟

ماهو المرمك؟

ماهو عداد البرامج؟

ما هو مسجل التجميع ومسجل التعليمه؟

كيف يتم قياس سعه المسجلات؟

الملفات النصيه التي يجب ان تسترجع

91.txt 92.txt 93.txt 94.txt 95.txt 96.txt 97.txt

98.txt 99.txt 100.txt 101.txt 102.txt 103.txt 104.txt

105.txt

## 8-هندسه البرمجيات

الملفات النصيه التي يجب ان تسترجع

112.txt 111.txt 110.txt 109.txt 108.txt 107.txt 106.txt

119.txt 118.txt 117.txt 116.txt 115.txt 114.txt 113.txt

120.txt

الاستفسار

ماهي هندسه البرمجيات؟

ماهو تاريخ هندسه البرمجيات؟

ما هي البرمجيات؟

لماذا هندسه البرمجيات؟

ما هي مراحل بناء النظام البرمجي؟

ما هي المهارات المطلوبه لتصيح مهندس

برمجيات؟

ماذا يقصد بمهندس برمجيات؟

الملفات النصيه التي يجب ان تسترجع

126.txt 125.txt 124.txt 123.txt 122.txt 121.txt

9 المودم

الاستفسار

ما هي مشاكل أجهزة المودم وإصلاحها؟

ما هي أجهزة المودم؟

كيف يتم تثبيت المودم او ازالته؟

ماهي الطريقة الصحيحة التي تقوي إرسال المودم

127.txt 128.txt 129.txt 130.txt 131.txt 132.txt

وتزيد من سرعة الإنترنت؟

ما هي انواع المودم؟

كيف تقاس سرعتها المودم؟

ماهي سرعات الاتصال التي تؤمنها المودمات؟

133.txt 134.txt 135.txt

كيف يعمل المودم؟

ماذا تعني كلمه مودم؟

10-الفأره

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هي الفأره ؟

ما هي اجزاء الفأره الرئيسيه؟

متى اخترعت الفأره؟

136.txt 137.txt 138.txt 139.txt 140.txt 141.txt 142.txt

ما هي انواع الفأره؟

كيف يتم توصيل الفأره اللاسلكيه؟

ما هي مميزات الفأره؟

143.txt 144.txt 145.txt 146.txt 147.txt 148.txt 149.txt

كيف يعمل فأره الحاسوب؟

ما هي فائده الفأره؟

ما هي الفأره الضوئيه؟

من هو مخترع الفاره؟

150.txt

## 11- وحدة المعالجه المركزيه

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هي انواع وحده المعالجه المركزيه؟

ماهو الناقل؟

ما هي وحده المعالجه المركزيه؟

ما هو وحده التحكم؟

ما هي العلاقه بين وحده المعالجه المركزيه

151.txt 152.txt 153.txt 154.txt 155.txt 156.txt 157.txt

والذاكره؟

ماهي وحده التنفيذ؟

كيف تعمل وحده المعالجه المركزيه؟

158.txt 159.txt 160.txt 161.txt 162.txt 163.txt 164.txt

ما هي اهميه وحده المعالجه المركزيه؟

مما يتكون وحده المعالجه المركزيه؟

ما هي انواع المعالجات؟

ما هي اجزاء المعالج الرئيسيه؟

165.txt

## 12- وحده الحساب والمنطق

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هي وحده الحساب والمنطق؟

ما هي انواع الذاكره؟

ما هو الحاسب الآلي؟

172.txt 171.txt 170.txt 169.txt 168.txt 167.txt 166.txt

مما يتكون الحاسب الآلي؟  
مما يتألف المعالج؟  
ما هي العمليات المنطقية؟  
مما يتكون وحدة الحساب والمنطق؟

179.txt 178.txt 177.txt 176.txt 175.txt 174.txt 173.txt

180.txt

### 13-تطبيقات الويب

الملفات النصية التي يجب ان تسترجع

الاستفسار

الفرق بين تطبيقات الويب ومواقع الويب؟

ماهو موقع الويب؟

ما هو تطبيق الويب؟

187.txt 186.txt 185.txt 184.txt 183.txt 182.txt 181.txt

وما هي امثله تطبيق الويب؟

ما هو اسلوب تطوير المواقع والتطبيقات؟

كيف تصبح مطور ويب؟

من المسئول عن تخزين البيانات؟

ما هو الفرق بين تطبيق الويب والموقع العادي؟

188.txt 189.txt 190.txt 191.txt 192.txt 193.txt 194.txt

ما هي المتاجر الالكترونيه؟

من هو العميل؟

من هو الخادم؟

195.txt

14- نظام التشغيل

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هو نظام التشغيل؟

ما هي مهام انظمه التشغيل؟

ما هو تاريخ نظم التشغيل؟

196.txt 197.txt 198.txt 199.txt 200.txt 201.txt 202.txt

ما هي انواع انظمه التشغيل؟

ما هي اشهر انظمه التشغيل؟

على ماذا يقوم هو نظام التشغيل؟

203.txt 204.txt 205.txt 206.txt 207.txt 208.txt 209.txt

ما هو اهداف نظام التشغيل؟

كيف يعمل نظام التشغيل؟

210.txt

## 15-القرص

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هو القرص المضغوط او المدمج؟

ما هو تاريخ وسائط التخزين؟

ما هو القرص الصلب؟

ماهو انواع الاقراص الصلبه؟

ما هي مكونات الاقراص الصلبه؟

ما هو القرص المرن؟

ما المقصود بالاقراص الاساسيه والديناميكيه؟

ما هو الاقراص البصريه؟

مما يتكون القرص البصري؟

ما هي تقنيه اتش دي دي؟

ما هي عيوب اتش دي دي؟

ما المقصود بتهيئه القرص الصلب؟

ما هو محرك القرص الضوئي او القرص

الليزري؟

217.txt 216.txt 215.txt 214.txt 213.txt 212.txt 211.txt

224.txt 223.txt 222.txt 221.txt 220.txt 219.txt 218.txt

225.txt

## 16\_الحاسوب

الملفات النصيه التي يجب ان تسترجع

الاستفسار

تعريف الحاسوب ؟

ما هي مكونات الحاسوب؟

خصائص ومميزات الحاسوب؟

ما هو تاريخ الحاسوب ؟

ما هو علم الحاسوب؟

مصطلحات مهمه في الحاسوب

231.txt 230.txt 229.txt 228.txt 227.txt 226.txt

237.txt 236.txt 235.txt 234.txt 233.txt 232.txt

كيف يعمل الحاسوب؟

240.txt 239.txt 238.txt

17\_الذاكرة:

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هي ذاكره الحاسوب؟

247.txt 246.txt 245.txt 244.txt 243.txt 242.txt 241.txt

ما هي انواع الذاكر المستخدمه في الحاسوب؟

ما هي مفهوم الذاكر الحاسوبيه؟

ما هي طريقه الوصول الى الموقع الذاكري؟

254.txt 253.txt 252.txt 251.txt 250.txt 249.txt 248.txt

ما هي ذواكر الوصول العشوائي؟

مصطلح الذاكره المقصود بها؟

255.txt

18\_انواع الحواسيب

الاستفسار

الملفات النصيه التي يجب ان تسترجع

ما هي انواع الحواسيب؟

262.txt 261.txt 260.txt 259.txt 258.txt 257.txt 256.txt

ما هي تصنيفات الحواسيب على حسب الوظيفة؟

تصنيفات الحواسيب على حسب الحجم؟

ما هي مميزات الحواسيب الكبيره؟

انواع الحاسبات الاليه؟

269.txt 268.txt 267.txt 266.txt 265.txt 264.txt 263.txt

ما هي اول نوع ظهر من انواع الجاسوب؟

270.txt

19\_وحداتي الادخال والايخراج في الحاسوب

الاستفسار

الملفات النصيه التي يجب ان تسترجع

تعريف وحداتي الادخال والايخراج في الحاسوب؟

277.txt 276.txt 275.txt 274.txt 273.txt 272.txt 271.txt

ما هي وظائف وحدات الادخال؟

ما هي وظائف وحدات الاخراج؟

ما هي اجهزه الادخال؟

ما هي اجهزه الاخراج؟

284.txt 283.txt 282.txt 281.txt 280.txt 279.txt 278.txt

ما هي وحدات المعاجه المركزيه؟

ما هي وحدات الادخال والايخراج معا؟

285.txt

20\_برمجيات الحاسوب

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هي برمجيات الحاسوب؟

292.txt 291.txt 290.txt 289.txt 288.txt 287.txt 286.txt

تنقسم الحاسوب الى ثلاثه برامج رئيسيه ما هي؟

كيف يستخدم الحاسوب في التعليم؟

ما هي استخدامات الحاسوب؟

299.txt 298.txt 297.txt 296.txt 295.txt 294.txt 293.txt

300.txt

21\_قواعد البيانات

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هي قواعد البيانات ؟

307.txt 306.txt 305.txt 304.txt 303.txt 302.txt 301.txt

ما هي مكونات وانواع قواعد البيانات؟

ما هي وظيفه قواعد البيانات؟

مفاهيم اساسيه في قواعد البيانات؟

اهميه قواعد البيانات؟

314.txt 313.txt 312.txt 311.txt 310.txt 309.txt 308.txt

مفهوم قواعد البيانات لنظم المعلومات؟

ما هي مشاكل الملفات في قواعد البيانات؟

ما هي قواعد البيانات الالكترونيه؟

315.txt

22\_نظم المعلومات المحوسب

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هي اسباب ظهور نظم المعلومات لمسانده

القرارات؟

322.txt 321.txt 320.txt 319.txt 318.txt 317.txt 316.txt

ما هي الفاظ البحث عن نظم المعلومات؟

ما هي انواع قطاعات نظم المعلومات؟

ما هي نظام المعلومات المحوسب؟

329.txt 328.txt 327.txt 326.txt 325.txt 324.txt 323.txt

ما هي نشاطات واجراءات نظم المعلومات

المحوسب؟

ما هي اهمية نظم المعلومات؟

330.txt

ما هي اهمية المدخلات في نظم المعلومات؟

23\_تقنيه المعلومات:

الاستفسار

الملفات النصيه التي يجب ان تسترجع

ما هي تقنيه المعلومات؟

337.txt 336.txt 335.txt 334.txt 333.txt 332.txt 331.txt

تعريف تكنولوجيا المعلومات؟

اهميه تكنولوجيا المعلومات؟

ما هي دور لمعلومات في الحياه؟

344.txt 343.txt 342.txt 341.txt 340.txt 339.txt 338.txt

ما هي اثر تكنولوجيا المعلومات؟

345.txt

24- نظام تخطيط موارد المؤسسه

الاستفسار

الملفات النصيه التي يجب ان تسترجع

ما هو نظام تخطيط موارد المؤسسه؟

كم من الوقت يستغرق بناء نظام تخطيط موارد

المؤسسه؟

352.txt 351.txt 350.txt 349.txt 348.txt 347.txt 346.txt

اهميه تخطيط موارد المؤسسه؟

أسباب قـدتؤدي إلى فشل نظام تخطيط  
مواردالمؤسسه؟

359.txt 358.txt 357.txt 356.txt 355.txt 354.txt 353.txt

ما هي برمجيات تخطيط مواردالمؤسسات؟  
من يحتاج الى تخطيط مواردالمؤسسه؟

360.txt

25\_تصميم وتحليل النظم:

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هو تحليل النظام؟

366.txt 365.txt 364.txt 363.txt 362.txt 361.txt 375.txt

ما هو تصميم النظام؟

ما هو النظام؟

من هو محلل النظم؟

373.txt 372.txt 371.txt 370.txt 369.txt 368.txt 367.txt

ما هي مدخلات مرحله التصميم؟

ما هي مخرجات عمليه التحليل؟

ما هي دوره حياة تتطور النظام؟

374.txt

26\_المحاكاة بالحاسوب:

الملفات النصيه التي يجب ان تسترجع

382.txt 381.txt 380.txt 379.txt 378.txt 377.txt 376.txt

389.txt 388.txt 387.txt 386.txt 385.txt 384.txt 383.txt

390.txt

الاستفسار

ما هو المحاكاة؟

ما هي النمذجه الحاسوبيه؟

ما هي خصائص نموج المحاكاة؟

ما مستويات النمذجه؟

ما هي اهداف المحاكاة؟

ما هي مميزات المحاكاة؟

تعريف المحاكاة بالحاسوب؟

27\_اداره المشاريع:

الملفات النصيه التي يجب ان تسترجع

397.txt 396.txt 395.txt 394.txt 393.txt 392.txt 391.txt

404.txt 403.txt 402.txt 401.txt 400.txt 399.txt 398.txt

405.txt

الاستفسار

ما هي مفهوم ادارة المشاريع؟

ماهي الخطوات الأساسية في إدارة المشاريع؟

ماهي المراحل الاساسية في ادارة المشاريع؟

اداره المشاريع اهميتهاوكيفيه تطبيقها؟

28\_اخلاقيات واداب استخدام الحاسوب:

الملفات النصيه التي يجب ان تسترجع

الاستفسار

ما هي اخلاقيات الحاسب الالى؟  
تنقسم اخلاقيات الحاسب الالى الى نوعين  
رئيسيين ما هي؟  
ما هي اقسام التعامل الاخلاقي؟  
ما هي أخلاقيات استخدام برامج الحاسوب؟  
ما هي أخلاقيات استخدام الحاسب بين الشخص  
وغيره؟

420.txt

## 29\_ جرائم الحاسوب

الاستفسار  
الملفات النصيه التي يجب ان تسترجع  
ما هي جرائم الحاسوب والانترنت؟  
ما هي أنواع جرائم الإنترنت؟  
ماهي الجريمه الإلكترونيه وما أنواعها؟  
تاريخ جرائم الانترنت؟  
ما هي ظاهرة جرائم الحاسوب؟

427.txt 426.txt 425.txt 424.txt 423.txt 422.txt 421.txt

434.txt 433.txt 432.txt 431.txt 430.txt 429.txt 428.txt

435.txt

الاستفسار

ما هي هندسه الكمبيوتر والحاسبات؟

ما هي تخصص هندسه انظمه الحاسبات؟

ما هي المجالات الوظيفيه لهندسه الحاسوب؟

ما هي اهتمامات هندسه الحاسوب؟

ما هو الفرق بين هندسه الحاسوب وهندسه

التحكم الالي؟

ما هو الفرق بين مهندس برمجيات ومهندس

حاسوب؟

الملفات النصيه التي يجب ان تسترجع

441.txt 440.txt 439.txt 438.txt 437.txt 436.txt 450.txt

448.txt 447.txt 446.txt 445.txt 444.txt 443.txt 442.txt

449.txt

الاستفسارات

الملفات النصيه التي يجب ان تسترجع

ما هو نظم استرجاع المعلومات ؟

456.txt 455.txt 454.txt 453.txt 452.txt 451.txt

ما هي نماذج استرجاع المعلومات ؟

ماهي مقاييس تقييم الأداء ؟

ما هي معدل الدقه ؟

ما هي معدل الاستدعاء ؟

462.txt 461.txt 460.txt 459.txt 458.txt 457.txt

ما الفرق الفرق بين نظم استرجاع المعلومات ونظم

المعلومات الأخرى ؟

متطلبات تطوير نظم استرجاع المعلومات ؟

465.txt 464.txt 463.txt

تطور ونشأة نظم الاسترجاع ؟

ما هي العوامل المؤثرة في عمليات استرجاع المعلومات ؟

ما هي أنواع أنظمة الاسترجاع ؟

كيف تبحث عن معلومة بكل سهولة ؟

## 32-الاقتصاد الكلي

الملفات النصيه التي يجب ان تسترجعها

الاستفسار

ما هو اقتصاد جزئي ؟

471.txt 470.txt 469.txt 468.txt 467.txt 466.txt

ما هو الاقتصاد الكلي ؟

ما هو تفسير العلاقة الطردية بين السعر وكمية العرض ؟

القوانين الأساسية للعرض والطلب ؟

الاقتصاد هو ؟

477.txt 476.txt 475.txt 474.txt 473.txt 472.txt

نشأة علم الاقتصاد ؟

انواع السلع ؟

ما هي المشكلة الاقتصادية ؟

480.txt 479.txt 478.txt

عناصر المشكلة الاقتصادية ؟

### 33- التجارة الإلكترونية

#### الاستفسار

عرف التجارة الإلكترونية؟

ما هو السوق الإلكتروني؟

ما هي الخطوات لبدء تجارك الإلكتروني؟

أهمية التجارة الإلكترونية؟

متطلبات التجارة الإلكترونية؟

تصنف التجارة الإلكترونية الى نوعين هما؟

ها و هيكل التجارة الإلكترونية؟

تطبيقات التجارة الإلكترونية؟

مشاكل التجارة الإلكترونية وحلولها؟

ما الفوائد التي يجنيها الزبائن من التجارة الإلكترونية؟

الملفات النصية التي يجب ان تسترجعها

486.txt 485.txt 484.txt 483.txt 482.txt 481.txt

492.txt 491.txt 490.txt 489.txt 488.txt 487.txt

495.txt 494.txt 493.txt

### 34- التسويق

#### الاستفسارات

ما هو التسويق؟

ما هي افضل طرق التسويق؟

هناك تعريفات كثيرة لعلم التسويق، أهمها؟

مما يتألف النشاط التسويقي؟

مميزات التسويق الإلكتروني؟

من هو العميل المستهدف؟

المستندات التي يجب ان تسترجعها

501.txt 500.txt 499.txt 498.txt 497.txt 496.txt

507.txt 506.txt 505.txt 504.txt 503.txt 502.txt

510.txt 509.txt 508.txt

ما هو التسويق المباشر؟  
نشأة و تطور مفهوم التسويق؟  
ما هي أهمية التسويق؟  
ما هي خطوات بحوث التسويق؟  
وظائف التسويق؟

المستندات التي يجب ان تسترجعها

516.txt 515.txt 514.txt 513.txt 512.txt 511.txt

522.txt 521.txt 520.txt 519.txt 518.txt 517.txt

525.txt 524.txt 523.txt

35- الذكاء الإصطناعي

الاستفسارات

ما هو الذكاء الإصطناعي؟  
أقسام الذكاء الإصطناعي؟  
ما هو علم المعرفة؟  
مشكلات الذكاء الاصطناعي؟  
تقييم الذكاء الاصطناعي؟  
تطبيقات الذكاء الاصطناعي؟  
تاريخ بدايه الذكاء الاصطناعي؟

المستندات التي يجب ان تسترجعها

36- إدارة علاقات العملاء

الاستفسارات

ما هو ادارة علاقات العملاء؟

531.txt 530.txt 529.txt 528.txt 527.txt 526.txt

أهم مزايا نظام إدارة علاقات العملاء؟  
ما هي أهميه إدارة علاقات العملاء؟  
تعريف إدارة علاقة الزبون؟  
تعريف إدارة العلاقات مع العملاء إلكترونياً؟

537.txt 536.txt 535.txt 534.txt 533.txt 532.txt

540.txt 539.txt 538.txt

المستندات التي يجب ان تسترجعها

546.txt 545.txt 544.txt 543.txt 542.txt 541.txt

552.txt 551.txt 550.txt 549.txt 548.txt 547.txt

555.txt 554.txt 553.txt

37- إدارة المعرفة

الاستفسارات

إدارة المعرفة ي قصد بها؟

أهمية إدارة المعرفة؟

مفهوم المعرفة؟

أهداف إدارة المعرفة؟

طرق معالجة إدارة المعرفة؟

### 38- أمن المعلومات

المستندات التي يجب ان تسترجعها

561.txt 560.txt 559.txt 558.txt 557.txt 556.txt

567.txt 566.txt 565.txt 564.txt 563.txt 562.txt

570.txt 569.txt 568.txt

الاستفسارات

ما هو أمن المعلومات؟

عناصر أمن المعلومات؟

ما الذي نريد ان نحميه؟

تتألف عملية إدارة المخاطر من؟

من هو الهاكر؟

تعريف فيروسات الكمبيوتر؟

التشفير هو؟

ما العمل ان تدقق أي من المخاطر رغم وسائل الحماية؟

### 39- بحوث العمليات

المستندات التي يجب ان تسترجعها

576.txt 575.txt 574.txt 573.txt 572.txt 571.txt

582.txt 581.txt 580.txt 579.txt 578.txt 577.txt

585.txt 584.txt 583.txt

الاستفسارات

ما هو بحوث العمليات؟

ما هي خطوات عملية صنع القرار؟

تطبيقات بحوث العمليات؟

أهم مواضيع بحوث العمليات؟

تاريخ بحوث العمليات؟

أهمية دراسة بحوث العمليات؟

السمات العامة لبحوث العمليات؟

### 40- ادارة سلسلة الامداد

المستندات التي يجب ان تسترجعها

الاستفسارات

تعريف سلسلة الامداد؟

591.txt 590.txt 589.txt 588.txt 587.txt 586.txt ما هو الفرق بين مفهوم ادارة سلسلة الامداد ومفهوم ال  
الخدمات اللوجستية؟

أهداف إدارة سلسلة الإمداد؟

597.txt 596.txt 595.txt 594.txt 593.txt 592.txt ما هي خطوات خلق سلسلة توريد فعّالة؟

لماذا نحتاج الى إدارة سلسلة التوريد؟

مفهوم سلسلة التوريد وإدارة سلسلة التوريد؟

600.txt 599.txt 598.txt