

نظام أرشفة وتحليل البيانات

لمركز الاستشارات والتنمية

مشروع تخرج متميز يهدف إلى تحويل الأرشفة اليدوية إلى نظام إلكتروني ذكي
متكامل يسهل حفظ البيانات والوصول إليها بسرعة وكفاءة عالية

تحت إشراف :

د/حسين الخضر

إعداد الطالب :

احمد عيسى جني
مازن عبده الحراري
محمد فوزي سايس
محمد نجيب زوبر

مشكلة البحث والتحديات الحالية

ضياع أو تلف الملفات

الملفات الورقية عرضة للتلف والضياع بسبب العوامل البيئية والتخزين غير الآمن

صعوبة البحث واسترجاع البيانات

الاعتماد على الأرشفة الورقية يجعل البحث عن المستندات عملية بطيئة وشاقة، مما يؤدي إلى فقدان الوقت الثمين

ضعف تحليل البيانات واتخاذ القرارات

صعوبة استخراج الرؤى والأنماط من البيانات الورقية يعيق عملية اتخاذ القرارات الإدارية

استهلاك الوقت والجهد

إدارة الملفات الورقية تتطلب موارد بشرية كبيرة وتکاليف تشغيلية عالية

أهداف المشروع الرئيسية

1

إنشاء نظام إلكتروني متكامل

بناء نظام رقمي شامل لأرشفة الملفات والمستندات بطريقة منظمة وأمنة

2

تمكين عملية البحث السريعة

توفير آليات بحث متقدمة تسمح بالوصول السريع والفعال إلى الوثائق المطلوبة

3

تطبيق تقنيات الذكاء الاصطناعي

استخدام خوارزميات متقدمة لتحليل البيانات النصية واستخراج الرؤى الذكية

4

تسهيل اتخاذ القرارات

توفير تقارير تحليلية شاملة تدعم عملية اتخاذ القرارات الإدارية والاستراتيجية

أهمية المشروع والفوائد المتوقعة



تقليل التكاليف المادية والزمنية

توفير موارد مالية وبشرية كبيرة بتقليل الحاجة لتخزين الورقي والموظفين المكلفين بالأرشفة اليدوية



رفع كفاءة إدارة البيانات

تحسين سرعة ودقة عمليات إدارة البيانات من خلال نظام إلكتروني منظم وألبي يقلل الأخطاء البشرية



دعم التحول الرقمي

المساهمة في رحلة التحول الرقمي للمؤسسات الأكاديمية والتنموية نحو بيئة عمل حديثة وفعالة



تحسين جودة الخدمات

تقديم خدمات أسرع وأكثر دقة للمستفيدين من خلال إمكانية الوصول السريع والموثوق للمستندات والبيانات

أدوات وتقنيات التطوير المستخدمة

معالجة البيانات



Python

لغة برمجة قوية لمعالجة البيانات وتطوير الخوارزميات
المقدمة

واجهة الأمامية



React JS

مكتبة JavaScript قوية لبناء واجهات مستخدم تفاعلية
وдинاميكية

واجهة الأمامية



HTML & CSS

لغات البناء الأساسية لتصميم واجهات الويب الحديثة
والمتغيرة

البيئة المحلية



XAMPP

حزمة متكاملة لتشغيل خوادم الويب والقواعد البيانات محلية

قاعدة البيانات



PostgreSQL

نظام إدارة قواعد بيانات علائقى قوى وموثوق وآمن

الذكاء الاصطناعي



Scikit-learn

مكتبة متخصصة لتطبيق خوارزميات تعلم الآلة والتصنيف
الألي

الخلفية النظرية والتقييات الأساسية



تعلم الآلة

Machine Learning

خوارزميات ذكية تتعلم من البيانات لتصنيف الوثائق واستخراج الأنماط والرؤى المهمة

تصنيف الوثائق تلقائياً

استخراج الأنماط والرؤى

تحسين الدقة المستمر



معالجة اللغة الطبيعية

NLP

تقنية تحليل النصوص واستخراج المعاني وال العلاقات بين الكلمات والمفاهيم المختلفة

تحليل النصوص العربية

استخراج الكلمات المفتاحية

فهم السياق والمعاني



التعرف الصوتي على الحروف

OCR

تقنية تحويل الصور الممسوحة ضوئياً والمستندات الورقية إلى نصوص رقمية قابلة للبحث والتعديل

تحويل الصور إلى نصوص

دعم اللغة العربية

معالجة الملفات المختلفة

الدراسات السابقة والأبحاث ذات الصلة

تصنيف الوثائق النصية باستخدام خوارزميات التعلم الآلي

Lee et al.

2019-2021

استخدمت هذه الدراسة خوارزميات **SVM** و **Logistic Regression** لتصنيف الوثائق النصية بدقة عالية

المهارات الرئيسية:

- مقارنة شاملة بين خوارزميات التصنيف
- تحقيق دقة تصنيف تتجاوز 95%
- تطبيق عملي على وثائق حقيقة

تحسين دقة أنظمة التعرف الصوتي على الحروف

Smith et al.

2020

تناولت هذه الدراسة تحسين دقة أنظمة **OCR** في استخراج النصوص من الصور والوثائق الموسوعة صوتيًا

المهارات الرئيسية:

- تطوير خوارزميات معالجة صور متقدمة
- تحسين معدل دقة الاستخراج بنسبة كبيرة
- دعم اللغات المتعددة بما فيها العربية

المعايير والمواصفات الدولية

ICA

International Council on Archives

ICA

معايير المجلس الدولي للأرشيف والممارسات العالمية

معايير الوصف الأرشيفي

تصنيف الوثائق الدولي

أفضل الممارسات العالمية

التعاون الدولي

NARA

National Archives and Records Administration

NARA

معايير الأرشفة الوطنية والسجلات الحكومية

حفظ الوثائق الحكومية

توثيق الإجراءات الإدارية

ضمان الحفاظ طويل الأجل

إناحة الوصول العام

ISO

International Organization for Standardization

ISO

معايير إدارة جودة البيانات والمعلومات

ضمان جودة البيانات

توحيد معايير التخزين

إدارة دورة حياة البيانات

ضمان الامتثال والشفافية

تحليل النظام والمكونات الرئيسية

وحدة رفع الملفات

تسمح للمستخدمين برفع المستندات والصور بصيغ متعددة (PDF، JPG، PNG) إلى النظام بشكل آمن وسهل

1

استخراج النصوص باستخدام OCR

تحويل الصور والملفات الممسوحة ضوئياً إلى نصوص رقمية قابلة للبحث والمعالجة باستخدام تقنية التعرف الصوتي

2

تحليل النصوص باستخدام NLP

معالجة النصوص المستخرجة لاستخراج الكلمات المفتاحية والمعانى و العلاقات بين المفاهيم المختلفة

3

تصنيف الوثائق والتخزين

تصنيف الوثائق تلقائياً حسب فئات محددة وتخزينها في قاعدة البيانات مع الفهرسة الكاملة للبحث السريع

4

ملاحظة مهمة:

تتعاون جميع هذه المكونات معاً في نظام متكامل يوفر أتمتة كاملة لعملية الأرشفة والتحليل، مما يقلل التدخل اليدوي ويزيد من الكفاءة والدقة

i

المتطلبات الوظيفية للنظام

1 تسجيل دخول المستخدمين

نظام تسجيل دخول آمن يتحقق من هوية المستخدمين ويوفر مستويات مختلفة من الوصول حسب الأدوار

2 رفع المستندات بصيغ متعددة

إمكانية رفع ملفات PDF وصور JPG وPNG مع التحقق من صحة الملفات وحجمها

3 استخراج النصوص تلقائياً

معالجة تلقائية لملفات المرفوعة لاستخراج النصوص باستخدام تقنية OCR بدقة عالية

4 تحليل وتصنيف الوثائق

تصنيف تلقائي للوثائق إلى فئات محددة باستخدام خوارزميات تعلم الآلة المتقدمة

5 البحث حسب الكلمات المفتاحية

محرك بحث قوي يسمح بالبحث السريع عن الوثائق باستخدام كلمات مفتاحية متعددة

6 إدارة صلاحيات المستخدمين

نظام متقدم لإدارة الصلاحيات يحدد من يمكنه الوصول والتعديل والحذف للوثائق

المتطلبات غير الوظيفية

سرعة الاستجابة



استجابة سريعة مع أداء عالي حتى مع كميات كبيرة
وقت استجابة أقل من ثانية
معالجة متوازية للملفات
تحسين الأداء المستمر

سهولة الاستخدام



تصميم واجهة مستخدم بديهية وسهلة الاستخدام
واجهة عربية متكاملة
تصميم بديهي وواضح
توثيق شامل وسهل

قابلية التوسيع



نظام قابل للتوسيع يتكيف مع النمو المستقبلي
معمارية قابلة للتطور
دعم قواعد بيانات كبيرة
سهولة إضافة ميزات جديدة

أمان البيانات وسريتها



حماية شاملة لبيانات من الوصول غير المصرح
تشفيير البيانات المخزنة
تشفيير الاتصالات
مراقبة الوصول والتدقيق

مخرجات ومنتجات النظام

واجهة استخدام سهلة



واجهة تفاعلية حديثة تدعم اللغة العربية وتتوفر تجربة مستخدم ممتازة
تصميم سهل الاستخدام

دعم كامل للعربية
واجهة متجلوبة
توثيق شامل

تقارير تحليلية شاملة



تقارير مفصلة تلخص أنواع الوثائق والمحنتى والإحصائيات المهمة
إحصائيات الوثائق

توزيع الفئات
تحليل الكلمات المفتاحية
رؤى قابلة للتنفيذ

قاعدة بيانات منظمة



نظام تخزين مقدم يحفظ جميع الوثائق والمستندات بطريقة
منظمة وأمنة

فهرسة شاملة للوثائق
بحث سريع وفعال
نسخ احتياطية دورية
حماية البيانات الكاملة

التحديات التي واجهت الفريق

1

دقة النصوص المستخرجة من الصور (OCR)

كان التعامل مع دقة استخراج النصوص من الصور والملفات الممسوحة ضوئياً تحدياً كبيراً، خاصة مع الوثائق القديمة والصور منخفضة الجودة

الحل المتبوع:

تطبيق معالجة صور مقدمة قبل استخدام OCR متعددة، وتطبيق خوارزميات تصحيح الأخطاء الذكية لتحسين الدقة

2

تحسين أداء خوارزميات التصنيف

تحقيق دقة عالية في تصنیف الوثائق تلقائیًا كان يتطلب تدريباً مكثفاً للنماذج وضبط المعاملات بشكل دقيق

الحل المتبوع:

استخدام عدة خوارزميات (SVM، Random Forest، Logistic Regression، Ensemble Learning)، تطبيق تقنيات (Random Forest، Logistic Regression)، واجراء تحسين مستمر بناءً على البيانات الفعلية

3

دمج الواجهة الأمامية مع خوارزميات التحليل

ربط واجهة المستخدم (React JS) مع الخوارزميات المعقدة (Python) كان يتطلب معمارية قوية وتواصل فعال بين الأنظمة المختلفة

الحل المتبوع:

بناء API قوية باستخدام الطوابير للمعالجة غير المتزامنة، وتطبيق نظام تخزين مؤقت فعال لتحسين الأداء

الفوائد والتطورات المستقبلية

توسيع الخدمة



توسيع النظام ليخدم مؤسسات أكاديمية وتنموية أخرى بنفس
الفعالية

- نسخ قابلة للتخصيص
- دعم متعدد المؤسسات
- معايير قابلة للتكيف
- توسيع الشبكة والتأثير

نظام عبر الإنترنت



تطوير النظام ليعمل كتطبيق ويب كامل يمكن الوصول إليه
من أي مكان وأي وقت

- إمكانية الوصول من أي جهاز
- دعم العمل عن بعد
- تحديثات تلقائية وفورية
- نسخ احتياطية سحابية

الخاتمة

ساهم هذا المشروع في بناء نظام فعال ومتكملا للأرشفة والتحليل الذكي للبيانات، مما يحقق نقلة نوعية في إدارة المعلومات داخل مركز الاستشارات والتنمية



دعم القرارات
توفير تحليلات ذكية لدعم اتخاذ
القرارات



تحسين الكفاءة
تقليل الوقت والجهد في البحث
والاسترجاع



تحويل رقمي كامل
من الأرشفة الورقية إلى نظام
إلكتروني ذكي

يعد هذا المشروع خطوة مهمة وحاسمة نحو التحول الرقمي الكامل للمؤسسات الأكademية والتنموية، ويفتح آفاقاً جديدة لتطوير الخدمات والعمليات الإدارية