2الجمهوريـــــة العربيـــــة الســـــورية
المعهد العالي للعلوم التطبيقية والتكنولوجيا
قســـــم المعلوميـــــات
العام الدراسي 2026/2025

مشروع تخرج

أعد لنيل درجة الإجازة في هندسة البرمجيات والذكاء الصنعي

نظام ذكي لإدارة المجلات العلمية المحكمة ودعم القرار التحريري

An Intelligent Scientific Journal Management System

تقديم الطالبة

كريستينا خيمي

إشراف

د. عمار الجر

م. مروى الشريف

م. بيان الأحمد

09/08/2026

الفصل الأول

التعريف بالمشروع

يعرض هذا الفصل تعريف المشروع، موضحاً أهدافه ومتطلباته.

1.1 مقدمة

تُعدّ المجلات العلمية المحكَّمة إحدى الوسائل الأساسية لنشر المعرفة وتقييم النتاج البحثي، إذ تمرّ الأبحاث المقدَّمة إليها بسلسلة من الإجراءات التي تبدأ باستقبال البحث والتحقق من استيفائه شروط الإرسال، ثم إخضاعه للفحص والتحكيم العلمي واتخاذ القرار التحريري المناسب، وتنتهي بإعداد الأبحاث المقبولة ونشرها ضمن أعداد المجلة. ويشارك في هذه الإجراءات عدد من الأطراف، مثل المؤلفين والمحكّمين والمحررين ورؤساء التحرير، مما يستدعي وجود نظام يضبط مسؤولياتهم وينظّم انتقال البحث المقدم بين مراحل المعالجة المختلفة.

2.1 الهدف من المشروع

يهدف المشروع إلى تصميم وبناء نظام إلكتروني ذكي ومفتوح المصدر يوفّر بيئة متكاملة لإدارة المجلات العلمية المحكَّمة، ويدعم دورة حياة النشر العلمي منذ إرسال البحث، مروراً بالتحكيم وإدارة التعديلات والقرارات التحريرية، وصولاً إلى إعداد الأبحاث المقبولة ونشرها وأرشفتها. ويسعى النظام إلى تنظيم إجراءات العمل وتوحيدها، وتحديد مسؤوليات المشاركين فيها وفق أدوار وصلاحيات واضحة. مع مراعاة بناء النظام بصورة قابلة لإعادة الاستخدام والتخصيص، بحيث يمكن تكييفه مع احتياجات المؤسسات الأكاديمية المختلفة وبذلك يتيح للمجلات الجامعية تبنّي حل رقمي ذكي بسهولة.

كما يهدف المشروع إلى تعزيز عملية اتخاذ القرار في النشر العلمي من خلال دمج تقنيات الذكاء الاصطناعي وتحليل النصوص العلمية لتوفير ميزات ذكية تشمل الكشف المبدئي عن الانتحال العلمي، والتوصية بالمحكّمين الملائمين للبحث استناداً إلى موضوع البحث وخبراتهم، وتحليل موضوعات الأبحاث واتجاهات النشر، وتوفير مؤشرات تساعد هيئة التحرير على متابعة العمل وتحديد أولويات المعالجة. وتؤدي هذه الخصائص دوراً مسانداً، من دون أن تحلّ محل التقييم العلمي أو القرار البشري الذي يبقى من مسؤولية الأطراف المختصة.

4.1- المتطلبات الوظيفية للنظام

استناداً إلى دراسة دورة العمل في المجلات العلمية المحكّمة، حُدّدت المتطلبات الوظيفية للنظام كما يأتي:

1. يسمح النظام للمؤلف بإنشاء حساب، وتسجيل الدخول، وإدارة بيانات ملفه الشخصي.
2. يسمح النظام لمدير النظام بإنشاء حسابات لطاقم العمل الممثل بالهيئة ومنح المستخدم الدور المناسب.
3. يتيح النظام للمستخدمين الذين يملكون اصلاحيات اللازمة تخصيص هوية المجلة وبياناتها وسياساتها وصفحاتها التعريفية، وإنشاء الأقسام العلمية وإدارتها وتعيين المسؤولين عنها، وتخصيص الأدوار والصلاحيات.
4. يمكّن النظام المؤلف من تقديم مخطوطته وبياناتها الوصفية وبيانات المؤلفين المشاركين، وإرفاق النسختين الكاملة والمعمّاة منها.
5. يتيح النظام للمؤلف متابعة حالة البحث المقدم، والاطلاع على سجل النسخ وقراراتها التحريرية، وإرسال النسخ المعدّلة عند طلبها.
6. يتيح النظام لمدير القسم فحص الأبحاث المقدمة الجديدة وفق قائمة تحقق محددة، ثم إحالتها إلى التحكيم أو رفضها مكتبياً مع توضيح السبب.
7. يسمح النظام لمدير القسم تعيين محرر مسؤول عن كل بحث، وإعادة إسناده عند الحاجة، ومتابعة تقدمه خلال مراحل المعالجة.
8. يتيح النظام للمستخدمين المؤهلين التقدم للعمل كمحكّمين ضمن المجلة وإرسال المعلومات اللازمة للتقدم، كما يسمح للجهات التحريرية بمراجعة طلباتهم وإدارة اختصاصاتهم وأعبائهم التحكيمية.
9. يقترح النظام محكّمين مناسبين للمخطوطة اعتماداً على مدى توافق موضوعها وكلماتها المفتاحية مع مجالات خبرتهم، مع إبقاء القرار النهائي للمحرر.
10. يمكّن النظام محرر القسم من اختيار المحكّمين لبحث ما وإرسال الدعوات وتحديد المهل الزمنية ومتابعة القبول والاعتذار والتأخر واستبدال المحكّمين عند الحاجة.
11. يضمن النظام إخفاء هوية المؤلفين عن المحكّمين وهوية المحكّمين عن المؤلفين، ويتيح تقديم تقارير التحكيم والتوصيات مع حماية الملاحظات السرية الموجهة للمحرر.
12. يمكّن النظام المحرر من إصدار قرار بالقبول أو الرفض أو طلب تعديلات، ويدعم رفع النسخ المعدلة وإجراء عدة جولات تحكيمية مع الاحتفاظ بسجل كل جولة.

13. يفحص النظام المخطوطات المقدمة بشكل مبدأي للانتحال العلمي ويولّد تقريراً عن حالات التشابه المحتملة ومصادرها، ليستخدمه المحررون والمحكمون بوصفه أداة مساعدة.

14. يتيح النظام إعداد سجلات النشر للأبحاث المقبولة، واستكمال بياناتها، وتنظيمها ضمن أعداد المجلة، ثم نشرها وأرشفتها.

15. يتيح للقراء تصفح المقالات المنشورة مع دعم البحث والتحميل

16. يدعم النظام محركات البحث الأكاديمية وتصدير بياناتها بصيغ توثيق أكاديمية(Dublin Core, RIS, BibteX)

17. يدعم النظام أرشفة أعداد المجلة، وإتاحة البيانات الوصفية للمقالات المنشورة لخدمات الفهرسة والاكتشاف الأكاديمي وفق بروتوكول **OAI-PMH**.

18. يرسل النظام إشعارات عند الأحداث المهمة، مثل دعوات التحكيم، وإسناد المحررين، وطلبات التعديل، والقرارات التحريرية، واقتراب المهل، ونشر المقالات.

19. يوفر النظام لوحات متابعة تتضمن مؤشرات تحليلية تساعد هيئة التحرير على مراقبة العمل واتجاهات النشر، وتحليل الموضوعات والاتجاهات البحثية، وترتيب المخطوطات التي تحتاج إلى متابعة وفق عوامل واضحة.

20. يؤمن النظام نموذج ترتيب للأبحاث المقدمة يساعد هيئة التحرير على تحديد أولويات المعالجة.

5.1- المتطلبات غير الوظيفية للنظام

يجب أن يحقق النظام المتطلبات غير الوظيفية الآتية:

1. أن يكون النظام آمناً بحيث يسمح فقط للمستخدم بالوصول إلى أقسام النظام بحسب الصلاحيات المتاحة له.

2. يجب أن يوفر النظام واجهات بيانية واضحة وسهلة الاستخدام

3. يجب أن يكون النظام قابل للتوسع وإعادة الاستخدام بشكل مرن

4. ألا يتجاوز زمن الاستجابة ثانيتين في 95% من الطلبات التفاعلية الاعتيادية، عند تشغيل النظام بوجود 50 مستخدماً متزامناً، وذلك باستثناء عمليات رفع الملفات وخوارزميات التحليل الذكي.

5. أن يستجيب النظام لطلب بدء العمليات طويلة التنفيذ، مثل فحص الانتحال وتحليل الموضوعات، خلال ثانيتين كحد أقصى، على أن تُنفّذ هذه العمليات في الخلفية مع إظهار حالتها للمستخدم.

6. ألا تتجاوز نسبة الطلبات الفاشلة 1% أثناء اختبار حمل يتضمن 50 مستخدماً متزامناً لمدة 15 دقيقة.

7. ألا تقل تغطية الاختبارات الآلية عن 70% من الشيفرة الخلفية، وعن 80% في الوحدات المسؤولة عن المصادقة والصلاحيات وانتقالات حالة المخطوطات والتحكيم والنشر.

8. ألا يكشف النظام هوية المؤلف للمحكّم، أو هوية المحكّم للمؤلف، أو الملاحظات التحريرية السرية لأي طرف غير مخوّل، وأن تنجح جميع سيناريوهات اختبار التحكيم مزدوج التعمية بنسبة 100%.

9. أن تتضمن كل نتيجة صادرة عن الخصائص الذكية درجة أو تفسيراً أو دليلاً يساعد المستخدم على فهمها، وألا تتسبب أي نتيجة ذكية بصورة مستقلة في قبول المخطوطة أو رفضها أو تغيير حالتها التحريرية.

10. ألا يؤدي فشل أي عملية إلى حفظ بيانات جزئية أو غير متسقة، وأن تنجح جميع اختبارات التراجع عن المعاملات المتعلقة بالإرسال والإسناد والقرارات والنشر بنسبة 100%.

الفصل الرابع

الدراسة التحليلية

يوضح هذا الفصل عملية تحليل النظام ودراسة متطلباته

1.4 مقدمة

يعرض هذا الفصل الدراسة التحليلية للنظام من خلال تحديد حدوده نطاق النظام وفهم الأطراف المتفاعلة معه وتحليل سير العمل الذي يجب تمثيله، ثم نمذجة حالات الاستخدام ودورة حياة المخطوطة. ويركز التحليل على الوظائف والقواعد التحريرية المطلوبة لدعم دورة النشر من الإرسال حتى القرار والنشر، مع اعتبار الميزات الذكية أدوات مساندة لاتخاذ القرار.

2.4 حدود النظام وسياقه

يشمل **نطاق النظام** إدارة دورة النشر العلمي للمخطوطات منذ تقديمها وبياناتها الوصفية، مروراً بالفحص الأولي وإسناد المحرر والتحكيم وجولات التعديل واتخاذ القرار، وصولاً إلى نشر المقالات المقبولة وإتاحتها. كما يشمل إدارة المستخدمين والأدوار والأقسام، وتوفير وظائف مساندة للعمل التحريري مثل الكشف المبدئي عن الانتحال، وتوصية المحكّمين، وترتيب الأولويات، وتحليل المؤشرات والموضوعات البحثية.

أما **خارج حدود النظام** فتقع كتابة البحث وإجراء تجاربه والتحقق العلمي من صحة نتائجه، وكذلك اتخاذ الحكم العلمي بدلاً من المحررين والمحكّمين. كما تقتصر علاقة خدمات الفهرسة ومحركات البحث الخارجية على إتاحة المحتوى والبيانات المنشورة لها.

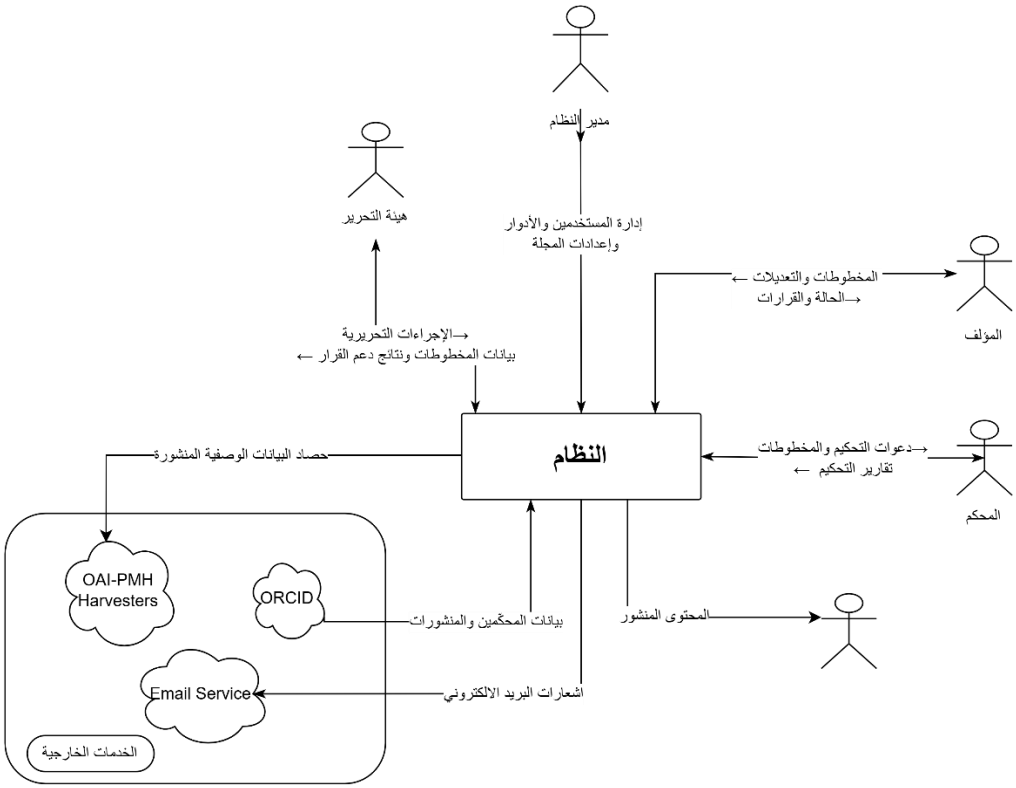


Figure 1: مخطط سياق النظام**system context diagram:**

3.4 تحليل الأطراف والأدوار في النظام

توزع مسؤوليات النظام على سبعة أدوار رئيسة، مع إمكان امتلاك المستخدم أكثر من دور عند استيفاء شروطه. وتختلف صلاحيات هذه الأدوار بحسب مشاركتها في التأليف أو التحكيم أو الإدارة التحريرية. وتتمثل الأطراف الرئيسة التي تتفاعل مع النظام في الآتي :

- **المؤلف**(**Author**) : يمثل الباحث الذي يقدم مخطوطته وبياناتها وملفاتها، ويتابع قراراتها، ويقدم النسخ المعدلة عند طلب التعديل.
- **المحكم**(**Reviewer**) : ستجيب لدعوات التحكيم، ويطّلع على النسخة المجهولة الهوية، ويقدم تقريره وتوصيته ضمن المهلة المحددة، وقد يشارك في تقييم النسخ المعدلة.
- **مدير القسم**(**Section Manager**) : يشرف على المخطوطات الواردة إلى قسمه، ويجري الفحص الأولي ويطّلع على تقرير الانتحال، ثم يرفض المخطوطة مكتبياً أو يسندها إلى محرر قسم.

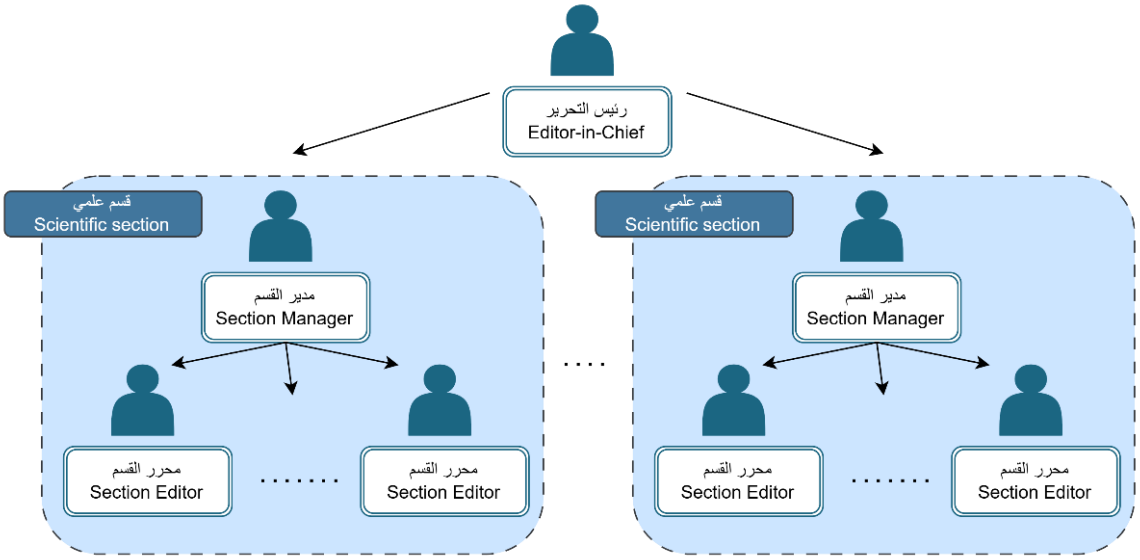
- **محرر القسم(Section Editor)** : يتابع المخطوطة بعد إسنادها إليه، ويتحقق من جاهزيتها للتحكيم، ويختار المحكّمين ويتابع الجولات، ثم يتخذ القرار التحريري ويحيل المقبول منها إلى النشر.
- **رئيس التحرير(Editor-in-Chief)** : يمثل أعلى سلطة تحريرية، ويشرف على نشاط المجلة عبر أقسامها ويتابع المؤشرات العامة وسير العمل.
- **مدير النظام(System Administrator)** : يدير المستخدمين والأدوار والإعدادات العامة، دون التدخل في الحكم العلمي أو القرارات التحريرية.
- **القارئ(Reader)** : يصل إلى المقالات والأعداد والمعلومات المنشورة للعامة، ولا يشارك في معالجة المخطوطات قبل نشرها.

العلاقة بين الأدوار التحريرية

يقتصر التسلسل الهرمي التحريري على رئيس التحرير ثم مديري الأقسام ثم محرري الأقسام، بينما يمثل المؤلف والمحكّم والقارئ أدواراً مشاركة خارج هذه الهرمية، ويعمل مدير النظام ضمن مسار إداري مستقل.

يمكن تمثيل هذه العلاقة في **مخطط هرمية الأدوار** بالشكل الآتي:

التسلسل الهرمي للأدوار التحريرية وعلاقة بقية الأطراف بالنظام



4.4 تحليل سير العمل العام للنشر العلمي

يمثل سير معالجة المخطوطة المسار الرئيس للنظام؛ إذ تنتقل من التقديم إلى الفحص الأولي والتحكيم والقرار، وقد تمر بجولات تعديل وإعادة تحكيم قبل انتهائها بالقبول والنشر أو بالرفض.

1.4.4 تقديم المخطوطة والفحص الأولي

يقدم المؤلف المخطوطة وبياناتها إلى القسم العلمي المناسب. يفحصها مدير القسم للتحقق من ملاءمتها واستيفائها المتطلبات الأساسية، مستفيداً من تقرير الكشف المبدئي عن الانتحال؛ ثم يقرر الرفض المكتبي أو استمرارها في المسار التحريري.

2.4.4 الإسناد والتحكيم العلمي

تُسند المخطوطة المقبولة أولياً إلى محرر قسم يتأكد من جاهزيتها للتحكيم، ثم يختار المحكّمين ويرسل الدعوات، مستفيداً من توصيات النظام. يبدأ التحكيم عند اكتمال العدد المطلوب من الموافقين، ويقدم كل محكّم تقريره وتوصيته ضمن المهلة المحددة.

3.4.4 القرار وجولات التعديل

بعد اكتمال التقارير يتخذ محرر القسم قراراً بالقبول أو الرفض أو طلب تعديلات طفيفة أو جوهرية. وعند طلب التعديل يقدم المؤلف نسخة جديدة مع الرد على الملاحظات، ثم تعاد إلى التحكيم، ويمكن تكرار هذه الدورة حتى الوصول إلى قرار نهائي.

4.4.4 القبول والانتقال إلى النشر

عند القبول تنتهي دورة التقييم وتنتقل المخطوطة إلى مسار إعداد المقالة للنشر ضمن أعداد المجلة وإتاحتها؛ أما المخطوطة المرفوضة فتنتهي معالجتها التحريرية.

5.4 تحليل حالات الاستخدام

جرى نمذجة النظام بمخططات حالات الاستخدام التالية

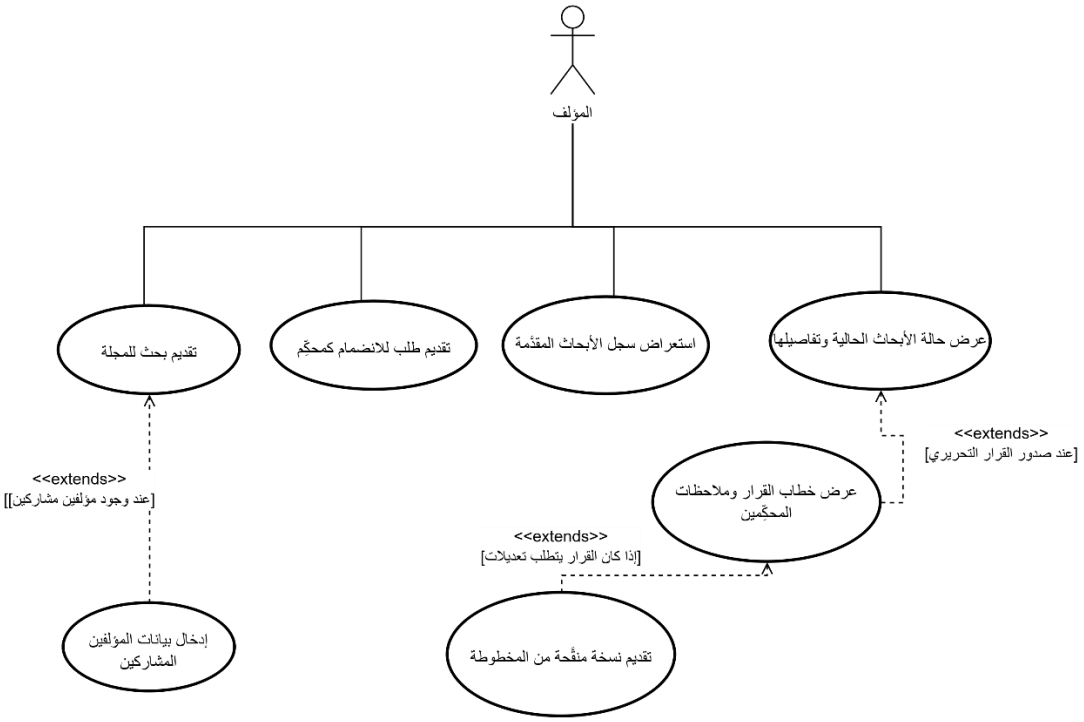
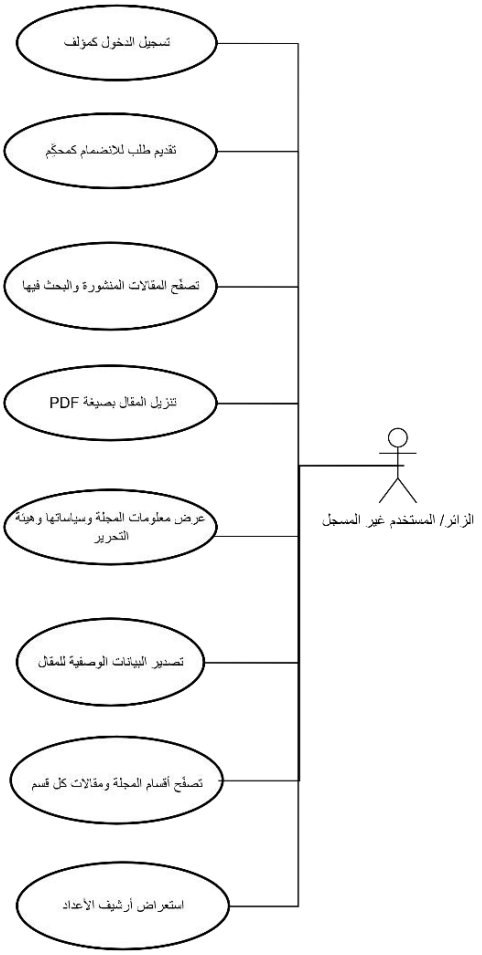


Figure 2 مخطط حالات استخدام المؤلف

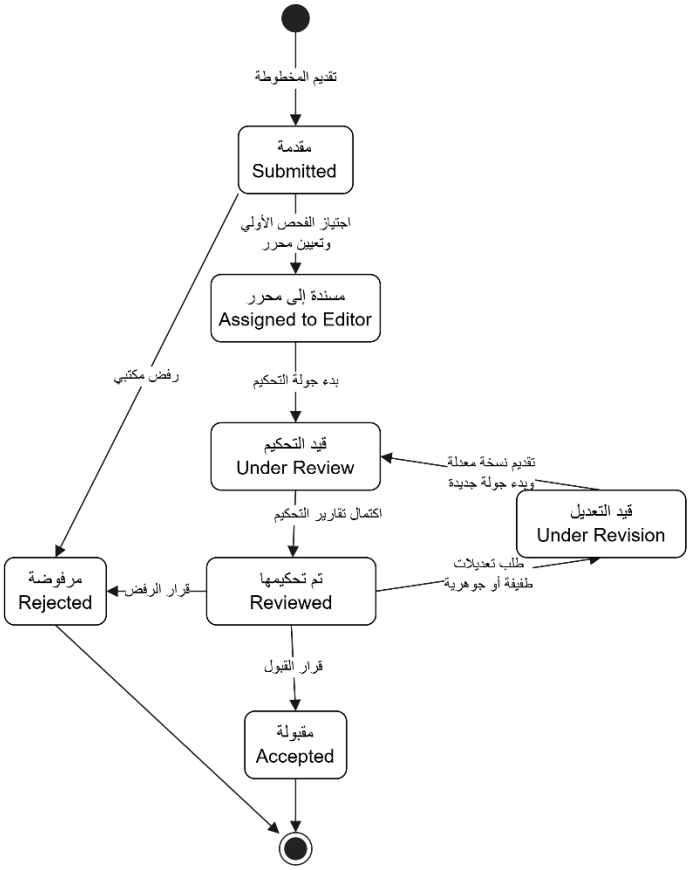
Figure 3 مخطط حالات استخدام المستخدم الغير مسجل



6.4 تحليل دورة حياة المخطوطة

تعبّر حالة المخطوطة عن موقعها في المسار التحريري. تبدأ «مقدَّمة»، ثم تنتقل بعد الفحص الأولي إلى «مسندة» فـ«قيد التحكيم»، وتصبح «تم تحكيمها» عند اكتمال تقارير الجولة. بعدها تُقبل أو تُرفض أو تنتقل إلى «قيد التعديل»؛ وعند تقديم نسخة معدلة تعود إلى «قيد التحكيم». وتمثل حالتا القبول والرفض نهاية الدورة التحريرية.

يوضح مخطط آلة الحالات الانتقالات السابقة وحلقة التعديل وإعادة التحكيم.


تقديم المخطوطة
مقدمة
Submitted
اجتياز الفحص الأولي
وتعيين محرر
مسندة إلى محرر
Assigned to Editor
رفض مكتبي
بدء جولة التحكيم
قيد التحكيم
Under Review
تقديم نسخة معدلة
وبدء جولة جديدة
قيد التعديل
Under Revision
اكتمال تقارير التحكيم
مرفوضة
Rejected
قرار الرفض
تم تحكيمها
Reviewed
طلب تعديلات
طفيفة أو جوهرية
قرار القبول
مقبولة
Accepted

الفصل الخامس

المنهجية المقترحة

يعرض هذا الفصل المنهجية المقترحة ضمن المشروع

1.5 مقدمة

5.1 مقدمة

تتضمن دورة النشر العلمي مهاماً لا تكفي لمعالجتها القواعد الإجرائية الثابتة وحدها؛ إذ تتطلب فهماً لمحتوى المخطوطة، ومقارنته بمصادر أو خبرات علمية أخرى، وموازنة عدة مؤشرات تحريرية في الوقت نفسه. وانطلاقاً من ذلك، تقترح هذه المنهجية مجموعة من الأساليب التي تحول النصوص والبيانات الوصفية وسجل سير العمل إلى مخرجات قابلة للمراجعة، من غير أن تنقل سلطة القرار من الإنسان إلى الخوارزمية.

تغطي المنهجية أربع وظائف تؤدي أدواراً متكاملة. يبحث فحص التشابه عن أدلة نصية موضّعة تربط مقاطع من المخطوطة بمصادر مفهرسة. وتكشف نمذجة الموضوعات البنية الموضوعية المتكررة ضمن مخطوطات القسم، وتستخلص كلمات تساعد على تفسير العناقيد المكتشفة. وتستخدم توصية المحكّمين التمثيلات الدلالية والكلمات المفتاحية لترتيب المحكّمين المؤهلين وفق ملاءمتهم العلمية. أما ترتيب الأولويات، فيجمع مؤشرات الزمن والحالة ونقص المحكّمين والتأخر وجولات التعديل في درجة تفسيرية تسهم في تنظيم المتابعة.

وتخضع جميع هذه المخرجات لمبدأ دعم القرار البشري. فلا تمثل درجة تشابه مرتفعة إثباتاً للانتحال، ولا تؤدي توصية محكّم إلى تعيينه تلقائياً، ولا تعبّر درجة الأولوية عن جودة المخطوطة أو احتمال قبولها. وإنما تهدف المنهجية إلى تقليص مساحة البحث، وإظهار الأدلة والعوامل المؤثرة، وتوجيه انتباه هيئة التحرير إلى الحالات التي تستحق مراجعة أعمق.

5.2 الإطار العام للمنهجية المقترحة

تنطلق المنهجية من مصدرين رئيسين للمعلومات: المحتوى العلمي للمخطوطة وبياناتها الوصفية من جهة، والبيانات التشغيلية المتولدة أثناء سيرها التحريري من جهة أخرى. ويُستخدم المحتوى النصي في الفحص المبدئي للتشابه، ونمذجة الموضوعات، وتقدير الملاءمة العلمية للمحكّمين؛ في حين تُستخدم حالة المخطوطة ومدد الانتظار وتقدم إجراءات التحكيم في ترتيب أولوية المتابعة. وبذلك تتكامل الإشارات الدلالية والنصية مع الإشارات التشغيلية من دون خلط وظائفها أو استخدام إحداها بديلاً عن الأخرى.

وتختلف الطبيعة الخوارزمية للمكوّنات الأربعة تبعاً للمشكلة التي تعالجها. يجمع كشف التشابه بين استرجاع حتمي معجمي ودلالي وبين نموذج تحقق متعلّم. وتعتمد نمذجة الموضوعات تعلّماً غير خاضع للإشراف لاكتشاف العناقيد من البيانات نفسها. أما توصية المحكّمين فهي مسألة ترتيب هجين تستفيد من تمثيلات دلالية مولّدة بنموذج سابق التدريب ومن تغطية الكلمات المفتاحية، وليست نموذجاً مستقلاً لاتخاذ قرار التعيين. ويعتمد ترتيب الأولويات دالة إرشادية موزونة وقابلة للضبط، لا نموذجاً متعلّماً من قرارات القبول والرفض السابقة.

يُنتج كل مسار مخرجاً قابلاً للتفسير: أدلة نصية ومصادر في فحص التشابه، وموضوعات وكلمات ممثلة في نمذجة الموضوعات، ومكونات درجة واضحة في توصية المحكّمين، وعوامل مساهمة في درجة الأولوية. ويتيح هذا المبدأ مراجعة سبب ظهور النتيجة وتقدير مدى صلاحيتها قبل الاستفادة منها في القرار التحريري. أما البنية البرمجية التي تستضيف هذه المكوّنات وآليات تشغيلها وتخزين مخرجاتها، فتُعرض في فصل تصميم النظام.

[ملاحظة لإدراج شكل: مخطط تدفق يوضح الإطار العام للمنهجية المقترحة؛ يبدأ بالمخطوطة ومحتواها وبيانات سير العمل، ثم يتفرع إلى أربعة مسارات هي الكشف المبدئي عن التشابه، ونمذجة الموضوعات، وتوصية المحكّمين، وترتيب الأولويات، وينتهي بمخرجات تفسيرية تُقدَّم إلى هيئة التحرير بوصفها أدوات دعم قرار مع إبقاء القرار النهائي بشرياً.]

5.3 منهجية الكشف المبدئي عن التشابه المحتمل والانتحال العلمي

5.3.1 صياغة المشكلة ونطاق الكشف

تُصاغ المهمة على أنها كشف خارجي عن التشابه على مستوى المقاطع؛ أي البحث، ضمن مجموعة من المصادر المفهرسة، عن مقاطع يمكن أن تفسر أجزاءً محددة من مخطوطة مقدمة. ولا تُعامل الوثيقة كاملةً بوصفها عنصراً واحداً للتصنيف، لأن المخطوطة الطويلة قد تضم حالات تشابه متعددة مع أكثر من مصدر، ولأن القيمة التحريرية لا تتحقق بنسبة كلية مجردة، بل بتحديد موضع النص محل الفحص والمصدر المقابل له.

يركز النطاق الحالي على النصوص العربية المستخرجة من ملفات نصية أو من ملفات PDF قابلة لاستخراج النص، مع الحفاظ على الأرقام والمصطلحات الأجنبية الواردة داخل النص العربي. ويقتصر الهدف على كشف تشابه محتمل يستدعي المراجعة؛ إذ قد ينشأ التشابه عن اقتباس موثق، أو مصطلحات تخصصية ثابتة، أو وصف شائع لمنهج علمي. لذلك لا تُفسَّر المخرجات باعتبارها حكماً على النزاهة العلمية، وإنما بوصفها أدلة أولية ينبغي فحص سياقها وطريقة توثيقها.

5.3.2 البيانات المرجعية وتكوين أمثلة التعلّم

استُخدمت مجموعة البيانات **ExAraCorpusPAN2015** مرجعاً لتطوير منهجية الكشف. وتضم المجموعة وثائق مصدرية، ووثائق مشتبه بها ضمن سياق مجموعة البيانات، وملفات توصيف بصيغة XML تحدد لكل حالة معلّمة هوية المصدر، ونطاق النص في الوثيقة محل الفحص، والنطاق المقابل في الوثيقة المصدرية. وتسمح هذه البنية ببناء أمثلة وتقييمات على مستوى المواضع النصية، بدلاً من الاكتفاء بإسناد وسم واحد إلى الوثيقة كاملة.

بعد تقسيم الوثائق إلى مقاطع، يُعد الزوج المكوّن من مقطع في الوثيقة محل الفحص ومقطع في المصدر مثالاً موجباً إذا تحققت ثلاثة شروط معاً: أن يتقاطع المقطع الأول مع النطاق المعلّم في الوثيقة محل الفحص، وأن ينتمي المقطع الثاني إلى الوثيقة المصدرية المحددة في التوصيف، وأن يتقاطع مع النطاق المعلّم داخل تلك الوثيقة. أما الأزواج التي لا تحقق هذه الشروط فتُعامل بوصفها أمثلة سالبة، بما فيها الأزواج التي تصل إلى الوثيقة الصحيحة ولكن إلى موضع غير صحيح. ويجعل هذا التعريف مهمة التحقق متسقة مع الهدف التشغيلي المتمثل في تقديم دليل نصي موضّع يمكن للمحرر مراجعته.

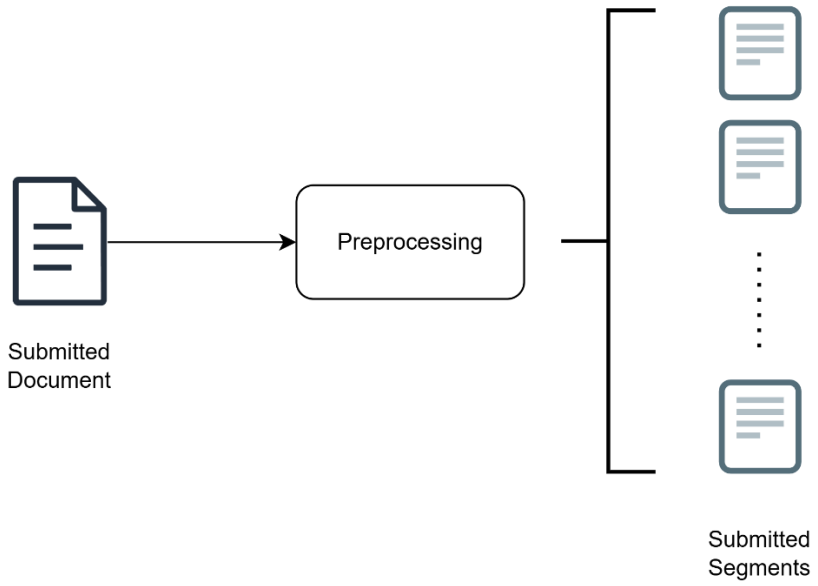
قُسمت بيانات التطوير إلى مجموعتي تدريب وتحقق بنسبة 80% و20% على مستوى الوثيقة محل الفحص كاملة، بحيث لا تتوزع مقاطع الوثيقة الواحدة بين المجموعتين. ويمنع هذا الإجراء تسرب أنماط نصية شديدة التقارب من الوثيقة نفسها إلى جانبي التقسيم. أما بيانات الاختبار فتُحجز عن التدريب واختيار نقطة الحفظ وعتبة القرار، كي تبقى مرحلة التقييم النهائي مستقلة عن قرارات الضبط.

5.3.3 استخراج النص والمعالجة المسبقة وتقسيمه

يبدأ الفحص باستخراج نص مرجعي موحّد من الوثيقة، ويُقصد به النسخة النصية التي تُنسب إليها إزاحات البداية والنهاية المستخدمة لاحقاً في التقرير. وتُطبّق معالجة عربية خفيفة تقلل الاختلافات الشكلية التي تضعف المطابقة من غير أن تغيّر البنية اللغوية للكلمات؛ فتُحذف علامات التشكيل والتطويل، وتُوحّد صور الألف، وتُحوّل الألف المقصورة إلى ياء، وتُسوّى المسافات وبعض علامات الترقيم، مع الإبقاء على الأرقام والمصطلحات الأجنبية، وعدم تطبيق تحويلات جذرية مثل تحويل التاء المربوطة إلى هاء.

تُحفظ لكل مقطع نسختان مترابطتان: النص الأصلي، ويُستخدم في التحقق والعرض، والنص المطبع، ويُستخدم في الاسترجاع المعجمي. ويُطبق مسار المعالجة نفسه على المخطوطة وعلى مجموعة المصادر، حتى لا تنشأ فروق مصطنعة بين جانبي المقارنة. كما تُحفظ إزاحات البداية والنهاية بالنسبة إلى النص المرجعي الأصلي، بما يتيح الرجوع إلى موضع الدليل وعرضه بدقة.

يعتمد التقسيم على حدود الفقرات ما أمكن، مع دمج الأجزاء القصيرة وتقسيم الأجزاء الطويلة. وتستهدف الإعدادات المعتمدة مقاطع متوسطة الحجم تقارب 800 محرف، ضمن حدود تقريبية بين 250 و1200 محرف، مع استخدام نافذة منزلقة على الكلمات عند تعذر التقسيم الطبيعي. ويوازن هذا الاختيار بين توفير سياق كافٍ لاستخلاص الدلالة وبين الحد من تشتت إشارة التشابه داخل مقطع طويل جداً.



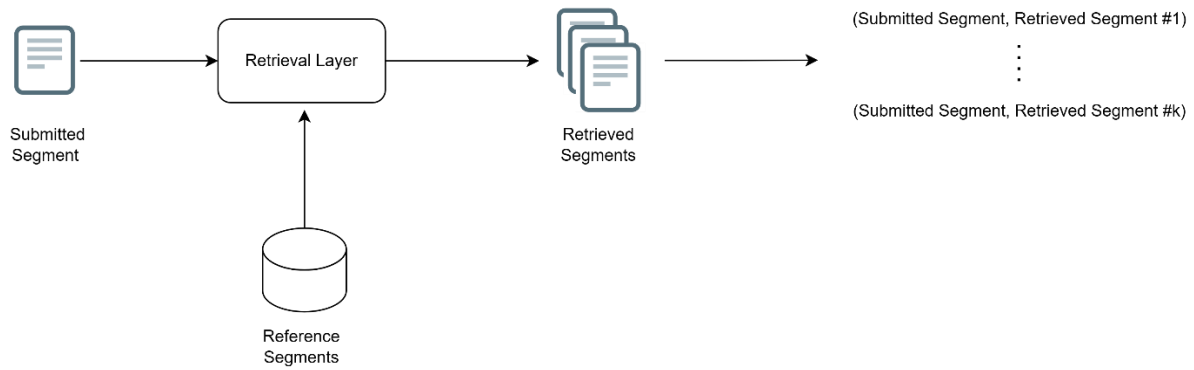
5.3.4 الاسترجاع الهجين للمقاطع المرشحة

تُعالج مجموعة المصادر مسبقاً بالطريقة نفسها، ثم تُمثّل في مسارين متكاملين. يستخدم المسار المعجمي خوارزمية **BM25** لاسترجاع حالات النسخ المباشر أو التعديل المحدود التي تحتفظ بنسبة مرتفعة من المفردات الأصلية. أما المسار الدلالي، فيستخدم نموذج **multilingual-E5** السابق التدريب لتوليد تمثيلات متجهية للمقاطع، ثم يقيس قربها الدلالي، بما يسمح باسترجاع نصوص متقاربة في المعنى حتى عند تغير جزء من الصياغة. ويُستخدم النموذج الدلالي في الاستدلال كما هو، من دون تدريب إضافي ضمن المشروع.

لكل مقطع من المخطوطة، يسترجع كل مسار قائمة مرشحين مستقلة. ثم تُوحّد القائمتان وتُزال الأزواج المكررة مع الاحتفاظ بمصدر الدليل المعجمي أو الدلالي أو كليهما. ولزوج مرشح يتألف من مقطع في المخطوطة ومقطع في المصدر، تُعرّف درجة الدمج المرجعية على النحو الآتي:

$$S_{\text{ret}}(q_i, s_j) = \max\{S_{\text{lex}}(q_i, s_j), S_{\text{sem}}(q_i, s_j)\},$$

حيث يمثل الحد الأول الدرجة النهائية للمسار المعجمي، ويمثل الحد الثاني الدرجة النهائية للمسار الدلالي، بينما تمثل القيمة العظمى بينهما درجة الدمج. ولا تُعد هذه الدرجة قراراً بوجود انتحال؛ فوظيفة الاسترجاع هي تحقيق استدعاء مرتفع وتقليص عدد الأزواج التي تنتقل إلى نموذج التحقق الأعلى كلفة.



5.3.5 التحقق العميق من الأزواج المرشحة

يُمرر كل زوج مرشح إلى نموذج **Siamese** مبني على مشفّر **AraT5v2** مشترك الأوزان. ويستقبل الفرع الأول مقطع المخطوطة، بينما يستقبل الفرع الثاني مقطع المصدر. ويحوّل المشفّر كل نص إلى سلسلة من التمثيلات السياقية ضمن الفضاء نفسه، ثم يُستخدم المتوسط المقنّع لاستبعاد رموز الحشو والحصول على تمثيل واحد لكل مقطع. فإذا مثّلت المصفوفة H التمثيلات السياقية للرموز، ومثّل المتجه a قناع الانتباه الثنائي المقابل لها، فإن تمثيلي المقطعين يُحسبان وفق:

$$\mathbf{h}_{q_i} = \frac{\sum_{t=1}^{T_q} a_t^{(q)}\, \mathbf{H}_t^{(q)}}{\sum_{t=1}^{T_q} a_t^{(q)}}, \qquad \mathbf{h}_{s_j} = \frac{\sum_{t=1}^{T_s} a_t^{(s)}\, \mathbf{H}_t^{(s)}}{\sum_{t=1}^{T_s} a_t^{(s)}}.$$

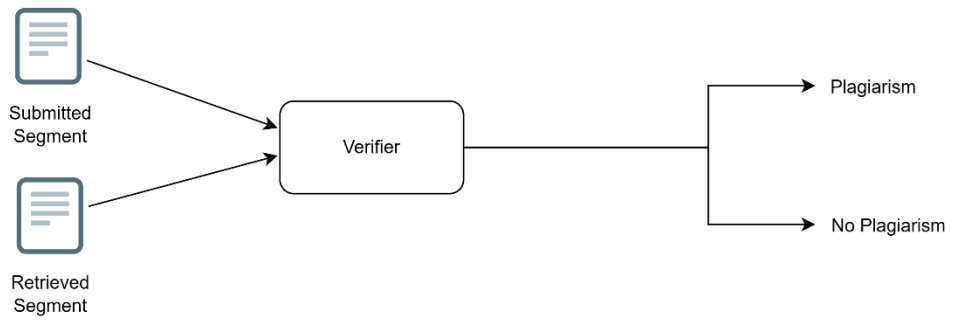
يمثل الرمزان Tq وTs عدد الرموز في مقطع المخطوطة ومقطع المصدر على الترتيب، بينما يمثل الرمزان hq وhs التمثيلين السياقيين الناتجين. وبعد ذلك تُبنى سمات مقارنة متناظرة تجمع الفرق المطلق والضرب العنصري:

$$\mathbf{z}_{ij} = \left[\left|\mathbf{h}_{q_i} - \mathbf{h}_{s_j}\right| ; \mathbf{h}_{q_i} \odot \mathbf{h}_{s_j}\right],$$

حيث يدل الرمز $\odot$ على الضرب عنصراً بعنصر، وتشير الفاصلة المنقوطة داخل القوسين إلى ربط متجهي السمات. ويُمرر متجه السمات الناتج إلى مصنف ثنائي متعدد الطبقات، ثم تُطبق دالة سيغمويد على خرجه لإنتاج احتمال التحقق:

$$p_{ij} = \sigma\left(\mathrm{MLP}(\mathbf{z}_{ij})\right), \qquad 0 \le p_{ij} \le 1.$$

يمثل الاحتمال الناتج مقدار ثقة النموذج بأن الزوج يطابق نمط الدليل النصي الذي تعلمه، ولا يمثل احتمالاً قانونياً أو أخلاقياً لوقوع الانتحال. ويجعل اشتراك الأوزان والسمات المتناظرة النتيجة معتمدة على العلاقة بين النصين لا على ترتيب إدخالهما. ويستخدم نموذج التحقق النص الأصلي للمقطعين، بينما يظل التطبيع موجهاً أساساً إلى الاسترجاع المعجمي، حفاظاً على الإشارات اللغوية التي قد تساعد النموذج على التمييز بين التشابه النصي الحقيقي والتقارب الموضوعي العام.



5.3.6 تدريب نموذج التحقق واختيار نقطة التشغيل

تتسم بيانات الأزواج بعدم توازن واضح، لأن الأمثلة السالبة تفوق الأمثلة الموجبة عدداً. لذلك تُستخدم في كل حقبة جميع الأمثلة الموجبة مع عينة سالبة بنسبة موجب إلى سالب قدرها 1:3، وتُدوّر الأمثلة السالبة بين الحقب لزيادة تنوعها. أما التحقق فيُجرى على التوزيع الطبيعي الكامل، كي لا تعكس مقاييس الضبط توزيعاً مصطنعاً أبسط من الحالة الفعلية.

يُضبط النموذج على مرحلتين. في المرحلة الأولى يُجمّد مشفّر AraT5 ويُدرَّب رأس التصنيف وحده. وفي المرحلة الثانية تُتاح آخر كتلتين من المشفّر للتعلّم بمعدل أصغر، مع إبقاء الطبقات الأخرى مجمدة. ويحقق هذا الضبط الجزئي توازناً بين الاستفادة من المعرفة اللغوية السابقة وتكييف الطبقات العليا مع مهمة مقارنة المقاطع، مع تقليل الكلفة الحاسوبية ومخاطر الإفراط في الملاءمة.

لنفترض أن مجموعة التدريب تضم N زوجاً، وأن y هو الوسم الحقيقي الثنائي لكل زوج، وأن p هو احتمال التحقق الذي ينتجه النموذج لذلك الزوج. تُستخدم خسارة الانتروبيا التقاطعية الثنائية:

$$\mathcal{L}_{\mathrm{BCE}} = -\frac{1}{N}\sum_{n=1}^{N}[y_n \log(p_n) + (1-y_n)\log(1-p_n)].$$

تُختار نقطة الحفظ الأفضل وفق **متوسط الدقة** (Average Precision)، لأنه يقيس جودة ترتيب الأمثلة الموجبة في حالة عدم التوازن من دون تثبيت عتبة مسبقاً. وبعد تثبيت النموذج، تُختار عتبة القرار على مجموعة التحقق بتعظيم مقياس F2:

$$F_2 = \frac{5\,\mathrm{Prec}\,\mathrm{Rec}}{4\,\mathrm{Prec} + \mathrm{Rec}},$$

حيث تمثل Prec الدقة وتمثل Rec الاستدعاء. ويعطي هذا المقياس وزناً أكبر للاستدعاء، وهو مناسب لمرحلة فحص أولي يُراد فيها تقليل احتمال إغفال الأدلة المحتملة، مع بقاء النتائج خاضعة للمراجعة البشرية. وتُثبت نقطة الحفظ والعتبة قبل التقييم على بيانات الاختبار. أما النتائج العددية والمقارنات بين التجارب فتُعرض في فصل الاختبارات والتقييم.

5.3.7 تحويل تنبؤات الأزواج إلى تقرير قابل للمراجعة

يُقبل الزوج دليلاً أولياً عندما يساوي احتمال التحقق الخاص به عتبة القرار أو يتجاوزها. ثم تُزال الأزواج المكررة، وتُدمج الأدلة المتجاورة عندما تشير إلى زوج المصدر نفسه وتكون نطاقاتها متقاربة في كل من المخطوطة والوثيقة المصدرية. ويحد هذا التجميع ثنائي النطاق من عرض الأدلة المجزأة بوصفها حالات مستقلة، مع الحفاظ على الارتباط بين موضع المخطوطة وموضع المصدر. وبعد ذلك تُنظم الأدلة حسب مقاطع المخطوطة ومصادرها كي يستطيع المحرر مراجعة النصين والدرجة المرتبطة بكل دليل.

ولحساب مقدار النص الذي تغطيه الأدلة المقبولة من غير احتساب التداخل أكثر من مرة، تُمثل نطاقات المقاطع المقبولة في النص المرجعي الموحّد للمخطوطة بفترات نصف مفتوحة، ويُمثل طول ذلك النص بالمحارف في مقام المعادلة الآتية:

$$C(m) = 100 \times \frac{\left|\bigcup_{k=1}^{n} I_k\right|}{L_m},$$

حيث يشير مقدار الاتحاد في البسط إلى الطول الكلي للنطاقات المقبولة بالمحارف بعد إزالة التداخل، ويرمز n إلى عدد هذه النطاقات. ولا تمثل نسبة التغطية الناتجة نسبة دقيقة للكلمات المنتحلة؛ بل مقدار النص الذي تغطيه الأدلة التي اجتازت عتبة التحقق. ويعرض التقرير، إلى جانب التغطية، عدد الأدلة والمصادر والمقاطع المعلّمة وأعلى احتمال تحقق، مع إبقاء كل دليل مرتبطاً بهوية المصدر ونصه الأصلي وإزاحاته وطريقة استرجاعه، وإظهار تنبيه صريح إلى ضرورة المراجعة البشرية.

5.3.8 حدود منهجية الكشف

ترتبط قدرة المنهجية مباشرةً بمجموعة المصادر المفهرسة؛ فالمصدر غير الموجود فيها لا يمكن استرجاعه مهما بلغت جودة نموذج التحقق. كما لا يشمل النطاق الحالي التعرف الضوئي على المحارف في ملفات PDF الممسوحة ضوئياً، ولا تحليل التشابه داخل الصور والمعادلات، وتقتصر صيغة التحقق المدروسة على المقاطع العربية. كذلك قد تُنتج النصوص المتقاربة موضوعياً أدلة كاذبة، وقد تفوت حالات أعيدت صياغتها جذرياً أو قُسمت بطريقة لا تحفظ الإشارة الكافية. وبناءً على ذلك، تُعد النتيجة أداة فحص مبدئي قابلة للتدقيق، لا حكماً نهائياً على وجود انتحال أو على النزاهة العلمية للمخطوطة.

5.4 منهجية نمذجة الموضوعات واستخراج الكلمات الممثلة

5.4.1 الهدف وتمثيل المخطوطات

تهدف نمذجة الموضوعات إلى اكتشاف البنية الموضوعية المتكررة داخل مخطوطات كل قسم من غير فرض قائمة تصنيفات ثابتة مسبقاً. ويُستخدم التمثيل المتجهي الدلالي لملخص كل مخطوطة لقياس القرب بين المخطوطات في الفضاء الدلالي، بينما تُستخدم نصوص الملخصات لتفسير العناقيد واستخراج الكلمات الممثلة لها. ويُشترط توافر حد أدنى قابل للضبط من المخطوطات، لأن العنقدة على عينات شديدة الصغر تكون غير مستقرة ولا تقدم موضوعات جديرة بالتفسير.

يتيح التمثيل الدلالي جمع مخطوطات تستخدم مفردات مختلفة لكنها تناقش مفاهيم متقاربة، وهو ما لا توفره طرائق عد الكلمات وحدها بالدرجة نفسها. وتظل كل نقطة في الفضاء المتجهي مرتبطة بالمخطوطة الأصلية، بحيث يكون الناتج إسناداً موضوعياً لكل مخطوطة، لا وصفاً منفصلاً للعناقيد فحسب.

5.4.2 اختزال الأبعاد والعنقدة الموضوعية

تُطبق منهجية **BERTopic** على التمثيلات الدلالية المحسوبة مسبقاً. وتبدأ باختزال الأبعاد باستخدام **UMAP** ومقياس جيب التمام، بهدف الاحتفاظ بالبنية المحلية الأكثر فائدة للعنقدة مع تقليل عدد الأبعاد والضجيج. ثم تنفذ BERTopic عنقدة كثافية تسمح باكتشاف عدد الموضوعات من البيانات نفسها بدلاً من تحديده مسبقاً، وهو أمر مناسب لاختلاف عدد الاتجاهات البحثية بين الأقسام والفترات الزمنية.

قد تُسند بعض المخطوطات إلى فئة الموضوع الشاذ، التي يرمز إليها BERTopic بالمعرّف «السالب واحد»، عندما لا تتوافر كثافة كافية لضمها إلى أحد العناقيد المكتشفة. وتُعامل هذه الحالات بوصفها قيماً شاذة بالنسبة إلى المجموعة الحالية، لا بوصفها مخطوطات بلا موضوع. كما يُعاد إجراء العنقدة عند تراكم عدد كافٍ من المخطوطات الجديدة، كي تعكس الموضوعات المحتوى المتاح حديثاً ولا تتحول النتيجة الأولية إلى تصنيف ثابت.

5.4.3 تمثيل الموضوع بالكلمات المميزة

بعد تكوين العناقيد، تُجمع نصوص كل عنقود بوصفها فئة نصية واحدة، ثم تستخدم BERTopic مقياس **c-TF-IDF** لتحديد الكلمات الأكثر تمييزاً لكل موضوع. وترتفع أهمية الكلمة عندما تتكرر داخل موضوع معين وتكون أقل شيوعاً في بقية الموضوعات. وتُختار الكلمات الأعلى وزناً لتكوين وصف موجز يساعد هيئة التحرير على تفسير مضمون العنقود وتسميته.

ولا تمثل هذه الكلمات ناتج خوارزمية مستقلة لاستخراج الكلمات المفتاحية لكل مخطوطة، بل وصفاً جماعياً للموضوع الذي تنتمي إليه. لذلك تُستخدم إشارةً تفسيرية إضافية، ويمكن ضمها إلى الكلمات التي أدخلها المؤلف عند توصية المحكّمين، من غير اعتبارها تصنيفاً إلزامياً أو حقيقة موضوعية ثابتة.

5.4.4 المخرجات والقيود

ينتج هذا المسار معرّف موضوع لكل مخطوطة منتمية إلى عنقود، ووصفاً موجزاً للموضوع، وقائمة بالكلمات الأعلى تمثيلاً له، مع تمييز الحالات الشاذة. ويمكن الاستفادة من هذه المخرجات في استكشاف الموضوعات النشطة وفهم توزيع المخطوطات داخل القسم، كما تقدم إشارة مساندة لتوصية المحكّمين.

ومع ذلك، تبقى النتائج استكشافية وتتأثر بحجم العينة وتوزيعها وإعدادات اختزال الأبعاد والعنقدة. وقد يتغير عدد الموضوعات أو تركيبها عند وصول مخطوطات جديدة، كما لا يعني اختلاف عنقودين بالضرورة وجود فصل علمي حاسم بينهما. لذلك لا تُستخدم المخرجات منفردة لاتخاذ قرار تحريري، بل تُفسر في ضوء محتوى المخطوطات والسياق العلمي للقسم.

5.5 منهجية توصية المحكّمين

5.5.1 بناء مجموعة المحكّمين المرشحين

تبدأ التوصية بتكوين مجموعة مرشحين مستوفية للقيود التحريرية قبل حساب أي درجة تشابه. ويُحصر البحث في المحكّمين الفعالين المرتبطين بقسم المخطوطة، ويُستبعد مؤلفها، ومن تكشف البيانات المتاحة عن تعارض في الهوية معه، ومن سبق توجيه دعوة إليهم في الجولة الحالية، ومن لديهم تكليف فعال للمخطوطة نفسها، ومن بلغوا الحد الأقصى المسموح للتكليفات الفعالة. وبهذا لا تستطيع درجة الملاءمة العلمية تجاوز شروط الأهلية أو تعارض المصالح؛ إذ تُحسب فقط داخل مجموعة مرشحين مقبولة تحريرياً.

5.5.2 تمثيل المخطوطة وخبرة المحكّم

يمثل ملخص المخطوطة بمتجه دلالي، لأنه يقدم وصفاً مركزاً لموضوع البحث وأهدافه. أما خبرة المحكّم فتمثل بمتجه مولّد من النصوص المتاحة في ملفه العلمي، وتشمل الكلمات المفتاحية والنبذة العلمية وعناوين المنشورات المتاحة عبر ORCID. وتُمنح الكلمات المفتاحية حضوراً واضحاً في وصف الخبرة لأنها تعبّر مباشرةً عن مجالات الاختصاص. ويُنشأ التمثيلان في الفضاء المتجهي نفسه، بما يجعل قياس التقارب الدلالي بين المخطوطة وخبرة المحكّم ممكناً.

وإلى جانب التمثيل الدلالي، تُبنى مجموعة كلمات المخطوطة من اتحاد كلمات المؤلف والكلمات الممثلة للموضوع المكتشف، بينما تتكون مجموعة كلمات المحكّم من كلمات خبرته بعد توحيد الصياغة وإزالة التكرار. ويجمع هذا التمثيل المزدوج بين التقارب في المعنى العام والتطابق الصريح للمصطلحات المتخصصة.

5.5.3 حساب درجة الملاءمة الهجينة

تُحسب درجة التشابه الدلالي للمخطوطة والمحكّم من تشابه جيب التمام بين متجهيهما، ثم تُحصر بين الصفر والواحد:

$$S_{\text{sem}}(m,r) = \text{clip}\left(\frac{\mathbf{v}_m^{\top}\mathbf{v}_r}{\| \mathbf{v}_m \|_2 \| \mathbf{v}_r \|_2}, 0{,}1\right),$$

حيث يمثل المتجه الأول ملخص المخطوطة، ويمثل المتجه الثاني وصف خبرة المحكّم، وترمز العلامة في المقام إلى المعيار الإقليدي. أما دالة الحصر فتُبقي القيمة بين الصفر والواحد. وتقيس النتيجة مقدار التقارب الدلالي العام بين موضوع المخطوطة ووصف خبرة المحكّم.

وعندما تكون مجموعة كلمات المخطوطة غير فارغة، تُحسب تغطية الكلمات المفتاحية وفق:

$$S_{\text{key}}(m,r) = \frac{|K_m \cap K_r|}{|K_m|}.$$

تقيس هذه الدرجة نسبة مفاهيم المخطوطة الصريحة التي يغطيها ملف المحكّم، ولذلك جُعل المقام عدد كلمات المخطوطة لا عدد كلمات المحكّم. ثم تُحسب درجة التوصية الهجينة:

$$S_{\text{rec}}(m,r) = \alpha\, S_{\text{sem}}(m,r) + \beta\, S_{\text{key}}(m,r), \qquad \alpha + \beta = 1.$$

عند توافر الإشارتين، يُستخدم الوزن الافتراضي 0.85 للتشابه الدلالي و0.15 لتغطية الكلمات، بحيث يكون التشابه الدلالي العامل الرئيس وتعمل تغطية الكلمات دليلاً مكملاً. وإذا تعذرت إحدى الإشارتين بسبب نقص البيانات، يُسند الوزن كاملاً إلى الإشارة المتاحة بدلاً من خفض الدرجة بصفر مصطنع؛ أما إذا غابت الإشارتان معاً فلا تتوافر أدلة كافية لتوصية المحكّم.

[ملاحظة لإدراج شكل: مخطط تدفق لمنهجية توصية المحكّمين يبدأ بتصفية المحكّمين غير المؤهلين وفق القسم والتعارض والدعوات السابقة وعبء العمل، ثم يعرض مسارين متوازيين: التشابه الدلالي بين ملخص المخطوطة ووصف خبرة المحكّم، وتغطية كلمات المخطوطة بكلمات المحكّم؛ وبعدهما دمج الدرجتين بالأوزان، ثم ترتيب القائمة وإظهار أسباب التوصية للمحرر.]

5.5.4 ترتيب النتائج وتفسيرها

يُرتب المحكّمون المؤهلون تنازلياً وفق درجة التوصية النهائية، ويُستخدم عدد التكليفات الفعالة لكسر التعادل لمصلحة المحكّم الأقل حملاً، ثم يُلجأ إلى ترتيب ثابت بالاسم عند استمرار التعادل. ولتعزيز قابلية التفسير، تُعرض درجة التشابه الدلالي، ودرجة تغطية الكلمات، والكلمات المتطابقة، وعبء العمل الحالي إلى جانب الدرجة النهائية.

ولا تنفذ القائمة عملية التعيين تلقائياً؛ فهي تقلص فضاء البحث أمام المحرر وتوضح أساس الترتيب. ويبقى القرار النهائي خاضعاً لعوامل قد لا تكون ممثلة بالكامل في البيانات، مثل التوفر الفعلي، والتعارضات غير المسجلة، ودقة التخصص، والخبرة الحديثة للمحكّم.

5.6 منهجية ترتيب أولوية معالجة المخطوطات

تُستخدم درجة الأولوية لترتيب المخطوطات بحسب حاجتها إلى المتابعة التحريرية، لا لتقدير جودتها العلمية أو احتمال قبولها. وتقوم المنهجية على خمسة عوامل قابلة للتفسير: مدة الانتظار منذ تقديم المخطوطة، وإلحاح الإجراء المطلوب في حالتها الحالية، ونقص عدد المحكّمين المقبولين عن العدد المطلوب، ووجود دعوات أو مراجعات متأخرة، وعدد جولات التعديل المنجزة. ويُطبّع كل عامل إلى درجة بين الصفر والمئة حتى يمكن دمج العوامل المختلفة على مقياس واحد.

لصياغة العوامل حسابياً، يرمز d إلى عدد أيام الانتظار، وD إلى المدة التي يبلغ عندها عامل الانتظار قيمته القصوى، وu إلى قيمة إلحاح الحالة وفق سياسة تحريرية معرّفة مسبقاً. ويرمز R إلى أعداد المحكّمين المطلوبين والموافقين، وN إلى مجموع الدعوات والمراجعات المتأخرة، وg إلى رقم جولة التعديل، وG إلى الحد الأقصى للجولات. وتُعرّف درجات العوامل كما يأتي:

$$
\begin{aligned}
s_1(m) &= 100\min\left(\frac{d_m}{D_{\max}}, 1\right),\\
s_2(m) &= u(m),\\
s_3(m) &= 100\,\frac{\max\left(R_{\mathrm{req}} - R_{\mathrm{acc}}(m), 0\right)}{R_{\mathrm{req}}},\\
s_4(m) &= 100\min\left(\frac{N_{\mathrm{late}}(m)}{R_{\mathrm{req}}}, 1\right),\\
s_5(m) &= 100\min\left(\frac{g_m}{G_{\max}}, 1\right).
\end{aligned}
$$

تمثل الدرجات الخمس، بالترتيب، أثر مدة الانتظار، وإلحاح الإجراء التحريري الحالي، ونقص تغطية التحكيم عندما تكون مطلوبة في الحالة الراهنة، وأثر الأعمال المتأخرة، وتقدم جولات التعديل. ويُضبط العامل الثالث إلى الصفر عندما لا تكون تغطية التحكيم مطلوبة في حالة المخطوطة الحالية.

إذا كانت أوزان العوامل الخمسة غير سالبة، على ألا تكون جميعها صفراً، فإن درجة أولوية المخطوطة تُحسب بالمتوسط الموزون:

$$P(m) = \frac{\sum_{i=1}^{5} w_i\ s_i(m)}{\sum_{i=1}^{5} w_i}.$$

تُستخدم افتراضياً الأوزان 25 و25 و20 و20 و10 للعوامل السابقة بالترتيب، مع إمكان تعديلها وفق سياسة المجلة. وتُرتب المخطوطات تنازلياً بحسب درجة الأولوية الناتجة، وتُستخدم أقدمية التقديم لكسر التعادل. كما تُعرض درجة كل عامل ومساهمته

في النتيجة النهائية، بما يمكّن المحرر من تفسير موضع المخطوطة في القائمة وتجاوز الترتيب الآلي عند وجود ظرف خاص لا تمثله العوامل المحددة.

5.7 خلاصة الفصل

قدّم هذا الفصل منهجية موحّدة لأربع وظائف ذكية تخدم مراحل مختلفة من العمل التحريري. يجمع فحص التشابه بين استرجاع هجين مرتفع الاستدعاء ونموذج تحقق متعلّم وتجميع للأدلة، وينتهي بتقرير موضّع يستلزم مراجعة بشرية. وتستخدم نمذجة الموضوعات تعلّماً غير خاضع للإشراف لاكتشاف البنية الموضوعية للمخطوطات ووصفها بكلمات ممثلة. أما توصية المحكّمين فتدمج التقارب الدلالي وتغطية الكلمات ضمن ترتيب يطبق بعد التحقق من أهلية المرشحين، في حين تنظّم درجة الأولوية المتابعة اعتماداً على عوامل تشغيلية موزونة وقابلة للتفسير.

وتتكامل هذه المناهج في تحويل المحتوى النصي وبيانات سير العمل إلى مؤشرات قابلة للفحص، من دون إسناد القرار النهائي إلى أي نموذج أو دالة حسابية. وبذلك تسهم المنهجية في رفع كفاءة المتابعة وتقليل الجهد اليدوي وتحسين اتساق الاستدلال التحريري، مع الحفاظ على مسؤولية هيئة التحرير وسلطتها في تفسير الأدلة واتخاذ القرار.

الفصل السادس

تصميم النظام

يعرض هذا الفصل القرارات التصميمية التي بني من خلالها النظام

1.6– مقدمة

يعرض هذا الفصل التصميم الهندسي لنظام إدارة المجلات العلمية الذكي، ويبين كيفية تنظيم مكوناته وتوزيع المسؤوليات بينها، وكيفية تمثيل بيانات النشر العلمي وضبط انتقالها عبر مراحل التقديم والتحكيم واتخاذ القرار والنشر. ويركز الفصل على بنية النظام والعلاقات وقواعد التصميم، لا على تفاصيل كتابة الشيفرة أو إعداد الأدوات أو نتائج التدريب والاختبار، إذ تعالج تلك الجوانب في فصولها المخصصة.

اعتمد التصميم على فصل واجهة الاستخدام عن الخادم الخلفي، وتنظيم الخادم الخلفي كتطبيق واحد معياري مقسم بحسب مجالات العمل. كما عوملت الخصائص الذكية بوصفها خدمات دعم قرار تتكامل مع سير العمل التحريري من خلال حدود واضحة، من دون منحها صلاحية اتخاذ قرار تحريري نهائي. وينطبق ذلك بوجه خاص على توصية المحكّمين وترتيب المخطوطات وكشف الانتحال؛ إذ تعرض هذه الخدمات درجات أو أدلة قابلة للفحص، بينما يبقى القرار للمستخدم المخول.

يصف الفصل التصميم المعتمد في النسخة الحالية من المشروع. لذلك لا يتضمن وظائف وردت ضمن المقترحات الأولية ولم تدخل في النظام المنفذ، مثل التحليل التنبؤي لقرار القبول أو تحليل شبكات الاستشهادات. كما يميز بين ما يعد جزءاً من النواة المشتركة وما يمكن تخصيصه عند تنصيب النظام لمجلة أخرى.

6.2 المعمارية العامة للنظام

6.2.1 النمط المعماري العام

يتبع النظام معمارية **العميل والخادم (Client–Server)**. تعمل واجهة الويب كعميل يتواصل مع الخادم الخلفي عبر واجهة برمجية تعتمد أسلوب REST، ويتولى الخادم التحقق من الهوية والصلاحيات وتطبيق قواعد العمل والوصول إلى البيانات والملفات والخدمات الذكية. يمنع هذا الفصل واجهة المستخدم من التعامل المباشر مع قاعدة البيانات أو مخزن الملفات أو النماذج الذكية، ويجعل الخادم الخلفي نقطة الضبط المركزية للعمليات الحساسة.

نظم الخادم الخلفي وفق **معمارية التطبيق الأحادي المعياري (Modular Monolith)**. فهو ينشر كتطبيق Django واحد، إلا أن داخله مقسم إلى وحدات نطاقية مستقلة نسبياً، مثل الحسابات والمخطوطات والتحكيم والنشر والنزاهة العلمية. اختير هذا الشكل لأنه يلائم ترابط دورة النشر العلمي؛ فالقرار التحريري يعتمد على المخطوطة وإصدارها وتقارير تحكيمها، والنشر يعتمد على قرار القبول. وفي الوقت نفسه يحافظ التقسيم المعياري على وضوح المسؤوليات ويحد من تحول المشروع إلى كتلة برمجية واحدة مترابطة بصورة يصعب صيانتها.

لا تعد عوامل Celery المنفصلة خدمات مصغرة بالمعنى المعماري؛ فهي تشغل مهاماً من قاعدة الشيفرة نفسها وتستخدم بيانات النظام نفسها. والغرض من فصلها تشغيلي، أي تنفيذ الأعمال الطويلة أو الدورية خارج دورة طلب HTTP، مع تخصيص عامل مستقل للعمليات الثقيلة الخاصة بكشف الانتحال.

أما التنظيم الداخلي، فيقارب بنية طبقية عملية تتكون من طبقة واجهات API، وطبقة خدمات تطبيقية لقواعد العمل المركبة، وطبقة نماذج ووصول إلى البيانات. ولا يفترض التصميم وجود طبقة نطاق معزولة بالكامل عن إطار Django؛ بل يستخدم طبقة الخدمات حيث تكون العملية أعقد من تحقق بسيط أو عملية قراءة وكتابة مباشرة.

6.2.2 القرارات المعمارية الرئيسة

القرار	تطبيقه في النظام	الغرض التصميمي
العميل والخادم	واجهة ويب مستقلة تتصل بخادم REST	فصل العرض عن منطق العمل والبيانات
التطبيق الأحادي المعياري	تطبيق خلفي واحد مقسم إلى وحدات نطاقية	تقليل التعقيد التشغيلي مع الحفاظ على تنظيم واضح
الاتصال عبر REST	موارد ومسارات موحدة بين الواجهة والخادم	توفير عقد تكامل واضح وقابل للتوثيق
طابور المهام والعوامل	مهام عامة ودورية وعامل متخصص للانتحال	إخراج الأعمال الطويلة من الطلبات التفاعلية
التخزين المخصص حسب نوع البيانات	بيانات المجال في PostgreSQL والملفات في MinIO	عدم خلط الملفات الثنائية بالبيانات العلائقية
دعم القرار البشري	مخرجات ذكية تفسيرية من دون قرارات آلية	الحفاظ على المسؤولية التحريرية والأخلاقية
الإعدادات بدلاً من تعديل الشيفرة	هوية المجلة ومحتواها وأقسامها وبعض السياسات قابلة للتهيئة	تسهيل إعادة استخدام النظام لدى جهة أخرى

ملاحظة لإدراج شكل (6-1): يوضع هنا مخطط معماري عام من نوع **C4 Container Diagram**. يعرض المستخدمين، وواجهة Next.js، وخادم Django REST، وPostgreSQL/pgvector، وMinIO، وRedis، وعامل Celery العام، وعامل كشف الانتحال، وCelery Beat، وخدمة البريد، وتكامل ORCID وأنظمة الحصاد عبر OAI-PMH. يجب أن يبين الشكل حدود النظام واتجاه الاتصال فقط، من دون إظهار وحدات Django الداخلية.

6.3 تصميم الوحدات والمكونات

6.3.1 وحدات الخادم الخلفي

قسم الخادم الخلفي بحسب مجالات النظام، بحيث تمتلك كل وحدة نماذجها وواجهاتها وخدماتها المتعلقة بمسؤولية محددة. ويوضح الجدول الآتي أهم الوحدات المنفذة:

الوحدة	المسؤوليات الرئيسة
الحسابات والأدوار accounts	المستخدمون والأدوار، ملفات خبرة المحكّمين، طلبات الانضمام إلى هيئة التحكيم، وبيانات ORCID
المجلة journals	الأقسام، عضويات محرري الأقسام، الأعداد، بيانات المجلة القابلة للتخصيص، صفحات المحتوى، والتحليلات التحريرية
المخطوطات submissions	بيانات المخطوطة، المؤلفون المشاركون، الإصدارات والملفات، ونتائج الموضوعات المشتقة
سير العمل workflow	الفرز الأولي، إسناد المحرر وإعادة الإسناد، انتقالات الحالات، وقوائم ترتيب الأولويات
التحكيم reviews	دعوات المحكّمين، الاستجابة للتكليف، التقارير، القرارات التحريرية، وإدارة التكليفات البديلة
النشر publishing	إنشاء سجل المقال من المخطوطة المقبولة، إعداد بياناته، ربطه بعدد، ونشره أو سحبه
الاكتشاف discovery	إتاحة البيانات الوصفية العامة ودعم بروتوكول OAI-PMH
النزاهة العلمية integrity	جدولة فحص الانتحال، حفظ حالته وتقريره، الصلاحيات، وربط النواة التحليلية بسير عمل المجلة
الإشعارات notifications	إرسال الرسائل الناتجة عن أحداث سير العمل والتذكيرات الدورية
الخدمات المشتركة core	التخزين، توليد التمثيلات المتجهية، وتوصية المحكّمين وخدمات مشتركة أخرى

تعتمد الوحدات باتجاه يعكس دورة العمل. فالمخطوطة ترتبط بالمؤلف والقسم، بينما تعتمد وحدتا سير العمل والتحكيم على المخطوطة وإصداراتها. وتعتمد وحدة النشر على مخطوطة وصلت إلى القبول وعلى الإصدار الذي استند إليه القرار، ثم تعتمد وحدة الاكتشاف على البيانات المنشورة فقط. أما وحدة النزاهة العلمية فترتبط بإصدار محدد وتستدعي نواة كشف الانتحال من خلال طبقة تكامل، من دون أن تصبح النواة مسؤولة عن حسابات المستخدمين أو حالات المخطوطات.

يقلل هذا الاتجاه من الاعتماد الدائري. ومن الأمثلة الواضحة تخزين نتيجة نمذجة الموضوعات ضمن SubmissionTopic في مجال المخطوطات، بدلاً من وضعها داخل وحدة المجلة، لأن النتيجة تخص مخطوطة بعينها حتى لو استخدمت لاحقاً في تحليلات القسم.

6.3.2 تصميم واجهة المستخدم

صممت الواجهة الأمامية وفق تنظيم قائم على المكونات والميزات. توجد عناصر عامة مشتركة للتنقل، وعرض الحالات، والنماذج والجداول، بينما تجمع كل ميزة صفحاتها ومكوناتها وأنواع بياناتها وعمليات الاتصال الخاصة بها. يحد ذلك من تكرار المكونات ويجعل تطوير مساحة عمل كل دور ممكناً من دون بناء تطبيق منفصل.

تقدم الواجهة مساحات عمل مختلفة بحسب الدور: مساحة للمؤلف لمتابعة مخطوطاته وإصداراتها، ومساحة للمحكّم للاستجابة إلى الدعوات وإرسال التقرير، ومساحة لمدير القسم للفرز والإسناد ومراجعة تقرير الانتحال، ومساحة لمحرر القسم لإدارة التحكيم والقرار، ومساحة لرئيس التحرير لإدارة الأعداد والاطلاع على التحليلات الشاملة، إضافة إلى البوابة العامة للقراء. هذا الفصل في العرض لا يعد بديلاً عن صلاحيات الخادم؛ فالواجهة تخفي الإجراءات غير المناسبة لتحسين تجربة الاستخدام، بينما يعيد الخادم التحقق من كل عملية.

ملاحظة لإدراج شكل (6–2): يوضع هنا **مخطط مكونات أو حزم (Component/Package Diagram)** يوضح وحدات الخادم السابقة واتجاهات اعتمادها. ينبغي أن يظهر plagiarism_core كنواة داخلية مستقلة نسبياً تتصل بوحدة integrity، لا كوحدة ذكاء اصطناعي عامة ولا كخدمة خارجية.

6.4 تصميم البيانات ونموذج المجال

6.4.1 المستخدمون والنطاق التنظيمي

يمثل المستخدم بكيان واحد يمكن أن يرتبط بأكثر من دور، لأن الشخص نفسه قد يكون مؤلفاً ومحكّماً في الوقت ذاته. غير أن الدور يمنح القدرة العامة فقط، بينما تحدد علاقات أخرى نطاق هذه القدرة. فمدير القسم يرتبط بالقسم الذي يديره، ويرتبط محرر القسم بعضوية فعالة في قسم معين، ولا تكفي حيازة الدور العام لإسناد مخطوطة من أي قسم إليه.

يحتفظ ملف المحكّم ببيانات الخبرة والكلمات المفتاحية والتمثيل المتجهي، ويرتبط بالأقسام التي يستطيع التحكيم فيها. وتستخدم هذه البيانات في الأهلية والتوصية، مع بقاء تعيين المحكّم إجراءً صريحاً يقوم به محرر القسم.

6.4.2 المخطوطة وإصداراتها

يفصل نموذج البيانات بين **المخطوطة (Submission)** بوصفها السجل المستمر للعمل التحريري، و**إصدار المخطوطة (SubmissionVersion)** بوصفه ملفاً وقراراً مرتبطين بمرحلة معينة. تحتفظ المخطوطة بالعنوان والملخص والكلمات المفتاحية واللغة والمؤلف والقسم والحالة الحالية والمحرر المسؤول، بينما يحتفظ كل إصدار برقم متزايد وملف المخطوطة الكامل والنسخة المجهلة وقرار ذلك الإصدار وخطاب القرار ورد المؤلف على المحكّمين.

يمنع هذا الفصل الكتابة فوق الملفات والقرارات القديمة عند طلب تعديل. فرفع نسخة معدلة ينشئ إصداراً جديداً، وتبقى جميع التكليفات والتقارير السابقة مرتبطة بالإصدار الذي راجعته فعلياً. وبذلك يمكن إعادة بناء التاريخ التحريري للمخطوطة من دون الاعتماد على سجل نصي غير منظم.

يحفظ SubmissionAssignment تاريخ إسناد المخطوطة إلى محرري الأقسام، في حين يشير حقل المحرر الحالي في المخطوطة إلى المسؤول النشط. ويجمع هذا التصميم بين سرعة الاستعلام عن المسؤول الحالي والمحافظة على أثر الإسناد وإعادة الإسناد وأسبابهما.

6.4.3 الفرز والتحكيم والقرار

يرتبط تقييم الفرز الأولي TriageAssessment بإصدار المخطوطة، ويحتفظ بنسخة قائمة التحقق والنتائج والملاحظات ومآل الفرز. ويسمح ذلك بتغيير قائمة التحقق في المستقبل من دون تغيير معنى التقييمات القديمة.

يرتبط تكليف المحكّم ReviewerAssignment مباشرةً بإصدار المخطوطة، وليس بكيان مستقل لجولة التحكيم. لذلك يمثل الإصدار الحد الفاصل بين الجولات في التصميم الحالي. ويحتفظ التكليف بحالته ومواعيد الاستجابة والتحكيم، ويمكن أن يشير إلى تكليف سابق حمل منه عند إنشاء إصدار معدل أو إلى تكليف ألغاه واستبدله. ويرتبط بكل تكليف تقرير تحكيم واحد يتضمن التوصية والتعليقات الموجهة للمؤلف والتعليقات السرية للمحرر.

هذا الاختيار يجعل العلاقة بين الملف الذي اطلع عليه المحكّم وتقريره علاقة صريحة. كما يمنع تكرار تكليف المحكّم نفسه للإصدار ذاته من خلال قيد فريد، ويطبق الحذف الوقائي على السجلات التي يعتمد عليها تاريخ العملية التحريرية.

6.4.4 النشر والبيانات المشتقة

يفصل التصميم بين المخطوطة المقبولة وبين المقال العام. ينشأ PublishedArticle كسجل في حالة المسودة من مخطوطة مقبولة، ويرتبط بالإصدار المصدر. وتنسخ إليه البيانات التي يجب أن تستقر عند النشر، مثل العنوان والملخص واللغة والكلمات المفتاحية وقائمة المؤلفين وترتيبهم، بدلاً من اشتقاق العرض العام دائماً من سجلات قد تتغير لاحقاً.

يرتبط المقال بعدد Issue عند إدخاله في خطة النشر. وللعدد دورة حياة مستقلة تشمل المسودة والنشر والأرشفة، مع تحديد العدد الحالي بصورة صريحة. ويساعد هذا الفصل رئيس التحرير على إدارة النشر المستمر والأرشيف من دون خلط حالة العدد بحالة المخطوطة التحريرية.

أما نتائج الخدمات الذكية فتعامل كبيانات مشتقة مرتبطة بأصولها. تخزن نتيجة نمذجة الموضوعات في سجل مرتبط بالمخطوطة، بينما تحفظ كل عملية كشف انتحال في PlagiarismScreening مرتبطة بإصدار محدد. ويسمح التصميم بوجود تاريخ من الفحوص للإصدار نفسه، مع منع وجود فحصين نشطين له في الوقت ذاته. ويمكن إعادة حساب نتائج مشتقة مثل التوصيات أو ترتيب الأولوية عند تغير البيانات الأصلية، بدلاً من معاملتها كحقائق ثابتة.

ملاحظة لإدراج شكل (6-3): يوضع هنا **مخطط كيانات وعلاقات منطقي (ERD)** يركز على الكيانات الجوهرية فقط: User وRole وSection وSubmission وSubmissionVersion وTriageAssessment وSubmissionAssignment وReviewerAssignment وReview وIssue وPublishedArticle وPublishedArticleAuthor وSubmissionTopic وPlagiarismScreening. يجب إظهار الكاردينالية والمفاتيح المرجعية المهمة، وعدم إضافة كيان ReviewRound لأنه غير موجود في التصميم الحالي.

6.5 تصميم سير العمل ودورات الحياة

لا تعامل حالة المخطوطة كحقل وصفي يمكن لأي واجهة تغييره مباشرةً. تحتفظ وحدة سير العمل بجدول صريح للانتقالات القانونية، وتستدعي الخدمات التطبيقية عملية الانتقال بعد التحقق من المستخدم والحالة الحالية والشروط السابقة. تقع حدود المعاملة والقفل على مستوى الخدمة التي تنفذ العملية، لأن عملية واحدة قد تعدل المخطوطة وإصدارها وتكليفاتها في الوقت نفسه.

يبدأ المسار بإنشاء المخطوطة وإصدارها الأول في حالة **مقدمة**. يجري مدير القسم الفرز الأولي باستخدام قائمة تحقق محفوظة، ثم يختار إما الرفض المكتبي أو السماح بالإسناد. وبعد الإسناد يصبح محرر القسم مسؤولاً عن اختيار المحكّمين وإدارة الدعوات. عند قبول العدد المطلوب من المحكّمين تنتهي الدعوات الزائدة، وعند اكتمال تقارير جميع المحكّمين المقبولين تصبح المخطوطة جاهزة لقرار المحرر.

ينتج القرار أحد ثلاثة مسارات رئيسة: القبول، أو الرفض، أو طلب تعديل بسيط أو جوهري. في حالة التعديل ينشئ المؤلف إصداراً جديداً، وتحفظ صلة التكليفات الجديدة بالتكليفات السابقة كي يبقى تاريخ جولات المراجعة قابلاً للتتبع. أما القبول فيسمح بإنشاء سجل المقال في مرحلة النشر، لكنه لا ينشره للعامة مباشرةً.

لا تستخدم دورة حياة واحدة لتمثيل جميع الكيانات. فحالة المخطوطة تصف تقدمها الإجمالي، وقرار الإصدار يثبت نتيجة جولة بعينها، وحالة تكليف المحكّم تصف استجابته، وحالة فحص الانتحال تصف عملاً خلفياً، وحالة المقال والعدد تصف النشر العام. الفصل بين هذه الحالات يمنع حالات متناقضة، مثل اعتبار المخطوطة منشورة لمجرد قبولها أو ربط تقرير تحكيم بإصدار آخر.

ملاحظة لإدراج شكل (6-4): يوضع هنا **مخطط دورات حياة مترابطة**، لا تكراراً لمخطط حالة المخطوطة في فصل التحليل. يعرض الشكل دورات المخطوطة، وقرار الإصدار، وتكليف المحكّم، وفحص الانتحال، والمقال المنشور، مع أسهم تبين أن طلب التعديل ينشئ إصداراً جديداً، وأن قبول الإصدار يتيح إنشاء مسودة المقال.

6.6 تصميم تكامل الخصائص الذكية

6.6.1 حدود خدمات دعم القرار

لا توجد في الخادم وحدة موحدة تحمل جميع وظائف الذكاء الاصطناعي؛ فقد وزعت كل وظيفة في المجال الأقرب إلى استخدامها. تشترك هذه الوظائف في مبدأين: يستدعيها الخادم نيابة عن الواجهة، وتعود بمخرجات قابلة للتفسير لا بقرار نهائي. ويوضح الجدول الآتي موضع كل خدمة:

الخدمة	المدخلات الرئيسة	المخرجات	موضع الاستخدام
توصية المحكّمين	تمثيل ملخص المخطوطة وكلماتها وملف خبرة المحكّم	قائمة مرتبة مع درجات وتفسير	مساحة عمل محرر القسم
نمذجة الموضوعات	ملخصات مخطوطات القسم وتمثيلاتها المتجهية	موضوعات وكلمات دالة وإسناد المخطوطات	التحليلات وتغذية التوصية
ترتيب الأولويات	حالة سير العمل والعمر والتكليفات والمواعيد	درجة مفسرة وعوامل مساهمة	قوائم العمل التحريرية ولوحة رئيس التحرير
التحليلات التحريرية	بيانات المخطوطات والتحكيم والنشر والموضوعات	مؤشرات وتوزيعات واتجاهات	لوحة رئيس التحرير
كشف الانتحال	ملف إصدار المخطوطة ومجموعة مصادر مفهرسة	تقرير أدلة على مستوى المقاطع	الفرز الأولي لدى مدير القسم

6.6.2 توصية المحكّمين

تبدأ التوصية بتكوين مجموعة مرشحين مؤهلين للقسم الذي تنتمي إليه المخطوطة. يستبعد النظام المؤلف، والمحكّمين غير النشطين، ومن سبق تكليفهم في الجولة الحالية، ومن لديهم تكليف نشط للمخطوطة نفسها، ومن تجاوزوا حد عبء العمل، إضافة إلى حالات تعارض الهوية الممكن اكتشافها من بيانات المؤلفين. بذلك لا تعوض درجة التشابه قواعد الأهلية، بل تطبق بعد تكوين مجموعة آمنة من المرشحين.

تعتمد درجة الملاءمة على مركبين: التشابه الدلالي بين تمثيل ملخص المخطوطة وتمثيل خبرة المحكّم، وتغطية الكلمات المشتركة بين كلمات المخطوطة والكلمات المستخرجة من موضوعها وملف المحكّم. يعيد النظام توزيع الأوزان عند غياب أحد مصدري الأدلة؛ فإذا لم يتوفر تمثيل متجهي يستخدم الكلمات فقط، وإذا غابت الكلمات يعتمد على الدليل الدلالي. وتعرض النتيجة سبب الترتيب، والكلمات المتطابقة، وعبء العمل الحالي، بدلاً من عرض رقم غير مفسر.

تبقى القائمة اقتراحاً لمحرر القسم. ولا يرسل النظام دعوة تلقائية إلى أعلى مرشح، لأن الملاءمة الموضوعية لا تغطي وحدها جميع التعارضات المهنية أو الاعتبارات التحريرية.

6.6.3 نمذجة الموضوعات والتحليلات

تنفذ نمذجة الموضوعات على مستوى القسم لأن تفسير المجموعة الموضوعية يكون أكثر معنى بين مخطوطات متقاربة في التخصص. تستخدم المهمة ملخصات المخطوطات وتمثيلاتها المتجهية المحسوبة مسبقاً، ثم تجري اختزال الأبعاد والتجميع واستخراج الكلمات المميزة للموضوع. تحفظ النتيجة في SubmissionTopic، وتعامل المخطوطات الشاذة صراحةً بوصفها غير مسندة إلى موضوع بدلاً من إجبارها على أقرب مجموعة.

التجميع عملية دورية لا تنفذ مع كل طلب. يتحقق مجدول المهام من وجود عدد كاف من المخطوطات الجديدة قبل إعادة تجميع القسم، لأن تشغيل نموذج موضوعي على عينات قليلة لا يعطي بنية مستقرة. وتستخدم الموضوعات في توزيع الاتجاهات البحثية وفي دعم الكلمات المستعملة في توصية المحكّمين، من دون أن تغير قسم المخطوطة أو حالتها آلياً.

تبني خدمة التحليلات لوحة شاملة لرئيس التحرير من البيانات التشغيلية المسجلة، وتشمل توزيع الحالات والمخطوطات والمنشورات عبر الزمن، ومعدلات القبول والرفض، ومدد القرار والتحكيم، وتوزيع الأقسام والموضوعات، والعمل المتأخر، وقائمة الأولويات. وتعرّف الخدمة معنى كل مدة أو مؤشر لتفادي تفسير أرقام متشابهة على نحو مختلف.

6.6.4 ترتيب أولويات المخطوطات

صمم ترتيب الأولويات كقاعدة موزونة شفافة، لا كنموذج يتنبأ بجودة البحث أو باحتمال قبوله. تجمع الدرجة عوامل مرتبطة بإدارة العمل: مدة الانتظار، والحاجة الحالية إلى إجراء تحريري، ونقص المحكّمين المقبولين، والدعوات أو التقارير المتأخرة، ورقم جولة التعديل. تطبّع كل قيمة إلى مجال موحد، ثم تضرب بوزن قابل للتهيئة على مستوى المجلة.

يعيد النظام مع الدرجة تفاصيل كل عامل ومساهمته وتفسيره، كي يستطيع رئيس التحرير معرفة سبب تقدم مخطوطة في القائمة. ولا تستخدم الدرجة لاتخاذ قرار قبول أو رفض، كما لا تدخل فيها خصائص قد تحول الترتيب الإداري إلى تقييم آلي للمحتوى العلمي.

6.6.5 التصميم العام لمنظومة كشف الانتحال

تعد منظومة كشف الانتحال جزءاً أصيلاً من المشروع، وليست استدعاءً لخدمة خارجية. صممت نواتها، ونفذت مراحلها، ودرب مكوّن التحقق فيها ضمن العمل على المشروع. وتعمل النماذج والفهارس محلياً داخل بيئة النظام عند توفير موارد التشغيل. أما تفاصيل إعداد بيانات التدريب والمعاملات والنتائج التجريبية فتترك لفصل التقييم، بينما يركز هذا القسم على البنية ومسار البيانات وحدود التكامل.

يقسم المكون إلى مستويين واضحين:

النواة التحليلية `plagiarism_core`: تضم إدخال الوثائق، والمعالجة المسبقة، وتوليد التمثيلات، والاسترجاع، والتخزين، والتحقق، وبناء التقرير، ومنسق خط المعالجة. لا تعتمد هذه النواة على نماذج حسابات المجلة أو واجهاتها.

طبقة التكامل `integrity`: تدير سجل الفحص وحالاته وصلاحياته ومهمته الخلفية، وتحصل على ملف الإصدار من التخزين الخاص، ثم تستدعي النواة وتحفظ التقرير وتعرضه ضمن الفرز الأولي.

يسمح هذا الفصل باختبار النواة وتشغيل أدوات فهرسة وتدريب خارج طلبات الويب، كما يمنع تسرب تفاصيل سير العمل التحريري إلى الخوارزميات. وفي الاتجاه المقابل لا تعرف وحدة المجلة تفاصيل BM25 أو E5 أو بنية نموذج التحقق؛ فهي تتعامل مع عقد تقرير محدد الإصدار.

> **ملاحظة لإدراج شكل (6-5):** يوضع هنا **مخطط حزم ومكونات لمنظومة كشف الانتحال**. يعرض حدود apps.integrity وplagiarism_core، ثم الوحدات الداخلية: ingestion وpreprocessing وstorage وembeddings وretrieval وverification وreporting وpipeline. ويظهر مخزن موارد النماذج والفهارس كمورد تشغيل محلي، لا كخدمة خارجية.

أ. إدخال الوثيقة والمعالجة المسبقة

يستقبل مسار المجلة في النسخة الحالية مخطوطة عربية بصيغة PDF نصي. تتحقق طبقة الإدخال من النوع والحجم وسلامة الملف والتشفير ووجود نص مضمّن قابل للاستخدام. إذا كان الملف ممسوحاً ضوئياً من دون نص فلا يدعي النظام تحليله، بل يعيد حالة توضح الحاجة إلى معالجة OCR، وهي وظيفة غير مدعومة حالياً.

يستخرج النص في صورة قانونية واحدة تحفظ فواصل الصفحات، ثم يمر النص عبر معالج مشترك يستخدم أيضاً عند فهرسة المصادر. توحد المعالجة بعض الاختلافات الكتابية العربية، مثل أشكال الألف والياء وإزالة التشكيل والتطويل، وتنظم علامات الترقيم والمسافات، مع الحفاظ على الأرقام والمصطلحات الإنكليزية. لا يستبدل النص الأصلي بالنص المطبع؛ إذ يحتفظ النظام بكليهما لاستخدام المطبع في المقارنة، والأصلي في التحقق وعرض الدليل.

بعد ذلك يقسم النص إلى مقاطع مناسبة للمقارنة. يبدأ التقسيم من الفقرات، ثم يدمج المقاطع القصيرة أو يقسم الطويلة، ويستخدم نافذة كلمات متداخلة عند الحاجة. يحمل كل مقطع رقمه ونصه الأصلي والمطبع ولغته وطوله وبدايته ونهايته في الوثيقة. وتستخدم إزاحات نصف مفتوحة من الشكل [البداية، النهاية)، بحيث يمكن إعادة استخراج النص المعروض في التقرير من الوثيقة الأصلية والتحقق من سلامة الإزاحات.

اختيرت المقارنة على مستوى المقاطع بدلاً من الوثيقة كاملة لثلاثة أسباب: تحديد موضع الدليل، والتعامل مع مخطوطة قد تجمع نصوصاً من مصادر متعددة، وتقليل طول المدخل إلى نماذج الاسترجاع والتحقق. كما يمنع المعالج المنسق انتقال أي مقطع ذي إزاحة لا تطابق نص الوثيقة إلى المراحل اللاحقة.

ب. فهرسة مصادر المقارنة

الفهرسة مسار مستقل يسبق فحص المخطوطة؛ فهي تنفذ عند إعداد مجموعة المصادر أو تحديثها. تمر وثائق المصدر بعملية الاستخراج والتطبيع والتقسيم نفسها، ثم يحفظ PostgresSourceRepository بيانات الوثيقة ومقاطعها في PostgreSQL. يتضمن كل مقطع النصين الأصلي والمطبع والإزاحات واللغة ومعرف المصدر، بينما يستخدم تجزئ النص لمنع إدخال المحتوى المكرر.

يبنى فوق المقاطع تمثيلان متكاملان. الأول فهرس معجمي BM25 يعتمد النص المطبع ويحفظ في ذاكرة مخبأة على القرص. والثاني تمثيل دلالي يولده نموذج E5 متعدد اللغات لكل مقطع مصدر. يدعم التصميم حفظ المتجهات والبحث التقريبي في pgvector، كما يدعم فهرساً دلالياً محلياً من مصفوفة متجهات وبيانات وصفية؛ والتهيئة الحالية تستخدم المسار المحلي افتراضياً مع بقاء PostgreSQL مستودعاً لمقاطع المصادر.

يجب أن تحمل قاعدة المصادر وذاكرة BM25 والفهرس الدلالي هوية المجموعة نفسها. لذلك يعامل نوع المصدر ومعرف الفهرس والبيان الوصفي للذاكرة كعقد اتساق واحد، وتستلزم إضافة المصادر أو حذفها إعادة بناء الأجزاء المتأثرة. يمنع ذلك مقارنة مخطوطة بفهرس معجمي لمجموعة وفهرس دلالي لمجموعة أخرى.

ج. الاسترجاع الهجين للمرشحين

لا يطبق نموذج التحقق على كل مقطع في مجموعة المصادر، لأن التكلفة تنمو سريعاً ولا حاجة إلى مقارنة معظم الأزواج. لذلك تسبق التحقق مرحلة استرجاع عالية الاستدعاء تضيق فضاء البحث.

يستعلم كل مقطع من المخطوطة في مسارين متوازيين منطقياً. يسترجع BM25 المقاطع التي تشترك في العبارات والمصطلحات مع النص المقدم، وهو مناسب للنسخ المباشر والتعديلات المحدودة. ويولد E5 تمثيلاً دلالياً للمقطع باستخدام صيغة الاستعلام، ثم يبحث عن أقرب تمثيلات المصادر، مما يساعد على استرجاع التشابه المعنوي الذي لا يظهر بتطابق حرفي واضح.

تدمج القائمتان في مرشحين موحدين، مع إزالة تكرار زوج المقطع المقدم ومقطع المصدر والاحتفاظ بدليل كل قناة ودرجتها ورتبتها. وتحول طبقة ProductionCandidateAdapter المرشح المسترجع إلى عقد الإدخال الذي يتوقعه نموذج التحقق، بما يشمل النصين الأصليين والإزاحات وبيانات الاسترجاع. تفصل هذه الطبقة بين تغير بنية الاسترجاع وثبات واجهة المتحقق.

عدد المرشحين المحتفظ بهم في كل قناة قابل للتهيئة. وقد استخدمت تجارب المشروع سياسات مختلفة لاختيار العمق والعتبة، لكن المقارنة التفصيلية بينها ونتائجها تعرض في فصل التقييم، ولا تعد جزءاً من وصف التصميم هنا.

د. نموذج التحقق من أزواج المقاطع

المرشح المسترجع يعني وجود تشابه معجمي أو دلالي، لكنه قد يكون نصاً في الموضوع نفسه لا حالة انتحال محتملة. لذلك يستخدم النظام نموذج تحقق ثنائي مدرباً ضمن المشروع لفصل الأزواج المحتملة عن المرشحين المرتبطين موضوعياً فقط.

بني المتحقق وفق معمارية سيامية (Siamese) تعتمد مشفراً مشتركاً من عائلة AraT5. يمر نص المخطوطة ونص المصدر في الفرعين نفسيهما، ثم يحسب تمثيل لكل نص باستخدام تجميع يراعي قناع الرموز حتى لا تدخل رموز الحشو في المتوسط. وتبنى خصائص مقارنة متناظرة من التمثيلين، مثل الفرق المطلق والضرب العنصري، ثم تمر عبر رأس تصنيف ثنائي ينتج احتمالاً لكون الزوج دليلاً محتملاً.

تحقق المعمارية المشتركة اتساق تمثيل الطرفين، بينما تحد الخصائص المتناظرة من اعتماد النتيجة على ترتيب النصين. ويحمّل العامل المخصص النموذج الأساس ونقطة الفحص المدربة وعتبة التشغيل مرة واحدة، ثم يعيد استخدامها في الدفعات التالية. ولا تعرض درجة المتحقق على أنها حكم قاطع؛ فهي بوابة لقبول دليل أولي إلى مرحلة التجميع.

هـ. تجميع الأدلة وتوليد التقرير

قد تنتج المقاطع المتداخلة عدة أزواج مقبولة تشير إلى الحالة نفسها. لذلك لا يحول كل زوج إلى مخالفة مستقلة، بل تجمع الأزواج حسب وثيقة المخطوطة ووثيقة المصدر، ثم تدمج الأزواج المتقاربة أو المتداخلة على جانبي المخطوطة والمصدر. ويسمى ذلك **التجميع ثنائي الامتداد** لأنه يشترط اتساق الموضعين، لا قرب المقاطع في المخطوطة وحدها.

يبني المكون بعد التجميع تقريراً ذا مخطط إصدار واضح هو plagiarism_report_v1. يتضمن التقرير بيانات الوثيقة وإصدار خط المعالجة وإعداداته، وملخصاً عاماً، وملخصات مرتبة للمصادر، وقائمة النتائج. وتحمل كل نتيجة النص المشتبه به وإزاحاته، وأدلة المصدر وإزاحاتها، واحتمال المتحقق، وبيانات الاسترجاع، وعلامة صريحة بأن النتيجة تتطلب مراجعة بشرية.

تحسب نسبة التغطية من اتحاد امتدادات المقاطع المعلّمة في المخطوطة، ولذلك تعني **نسبة النص المغطى بالأدلة المحتملة**، ولا تعني نسبة مؤكدة من الكلمات المنتحلة. كذلك فإن وجود نتيجة تجاوزت عتبة التشغيل لا يثبت سوء السلوك العلمي؛ بل يعني أن النظام وجد دليلاً يستحق فحص مدير القسم. هذه الدلالة جزء من التصميم الأخلاقي للتقرير وليست مجرد صياغة في الواجهة.

بعد إنشاء الإصدار العربي الأول وتثبيت معاملة التقديم، ينشئ النظام سجل فحص ويرسل مهمته إلى طابور plagiarism. ينزل العامل ملف المخطوطة الكامل من MinIO إلى مسار مؤقت، ويستدعي كاشفاً محملاً مرة واحدة في عملية العامل، ثم يتحقق من إصدار مخطط التقرير قبل حفظه. لا ترسل المخطوطة إلى مزود نماذج خارجي.

يمر سجل الفحص بالحالات: في الانتظار، وقيد التنفيذ، ومكتمل، وفاشل. ويحفظ معرف المهمة وهوية خط المعالجة ونقطة الفحص وملخص التقرير والتقرير الكامل أو رمز فشل آمن. يقيد التصميم كل إصدار بفحص نشط واحد، ويطالب العامل بالسجل تحت قفل قبل التنفيذ، ولا يثبت نتيجة النجاح أو الفشل إلا إذا بقيت ملكية المهمة وحالتها متوافقتين. تقلل هذه القواعد آثار إعادة تسليم الرسالة أو تشغيل المهمة بصورة متزامنة.

فشل الفحص لا يلغي تقديم المخطوطة ولا يغير حالتها التحريرية. يستطيع مدير القسم المسؤول عن القسم رؤية التاريخ والملخص والأدلة وإعادة طلب الفحص عند فشله، بينما لا تتاح الأدلة الكاملة لجميع المستخدمين. ويحافظ ذلك على سرية المواد وعلى الفصل بين مؤشر النزاهة والقرار التحريري.

ملاحظة لإدراج شكل (6-6): يوضع هنا مخطط تدفق ذو مسارين. يعرض المسار الأول **فهرسة المصادر مسبقاً:** مصادر المقارنة ← الاستخراج والمعالجة المشتركة ← التقسيم ← مستودع PostgreSQL ← فهرس BM25 وتمثيلات E5. ويعرض المسار الثاني **فحص المخطوطة:** إصدار من MinIO ← الاستخراج والتقسيم ← الاسترجاع الهجين ← دمج المرشحين ← Adapter ← متحقق AraT5 ← التجميع ثنائي الامتداد ← plagiarism_report_v1 ← مراجعة مدير القسم. يجب ألا يوجد سهم يغير حالة المخطوطة آلياً.

ملاحظة لإدراج صورة (6-7): يستحسن إدراج لقطة واحدة من نافذة تقرير الانتحال داخل مساحة الفرز الأولي، على أن تظهر الملخص وقائمة المصادر ودليلاً واحداً على مستوى المقاطع. تضيف هذه الصورة قيمة لأنها تبين كيف ترجم التصميم التفسيري للتقرير إلى مادة قابلة للمراجعة، ولا حاجة إلى لقطات إضافية لكل حالة من حالات الفحص.

6.7 تصميم الأمن والتحكم بالوصول

يعتمد التحكم بالوصول على ثلاثة عناصر مجتمعة: **دور المستخدم، ونطاقه التنظيمي، وعلاقته بالمورد.** فدور مدير القسم لا يمنح الوصول إلى جميع الأقسام، ودور محرر القسم لا يمنح التعامل مع كل المخطوطات، ودور المحكّم لا يكشف أي مخطوطة لمجرد وجوده في مجموعة المحكّمين.

تطبق القواعد الأساسية في الخادم الخلفي، ومن أمثلتها:

يصل المؤلف إلى مخطوطاته وإصداراتها وخطابات القرار الخاصة بها.

يدير مدير القسم المخطوطات الواردة إلى القسم الذي يديره، ويجري الفرز والإسناد ويطلع على تقرير الانتحال.

يصل محرر القسم إلى المخطوطات المسندة إليه، ويختار المحكّمين ويدير التقارير والقرارات.

لا يصل المحكّم إلى مواد التحكيم إلا من خلال تكليف يخصه، وتراعى حالته قبل إتاحة الإجراء أو الملف.

يصل رئيس التحرير إلى التحليلات الشاملة وإدارة الأعداد والنشر وفق صلاحياته.

تعرض البوابة العامة المقالات والأعداد المنشورة والبيانات الوصفية العامة فقط.

يدعم التصميم التحكيم المزدوج المجهل من خلال فصل الملف الكامل عن الملف المجهل داخل إصدار المخطوطة، ومن خلال تقديم بيانات مختلفة بحسب الدور. فلا تعرض هوية المؤلف للمحكّم ضمن حزمة التحكيم، ولا تكشف هوية المحكّم للمؤلف، كما تبقى تعليقات المحكّم السرية للمحرر منفصلة عن التعليقات المرسلة إلى المؤلف.

تحفظ ملفات المخطوطات وتقاريرها في مخزن كائنات خاص، وتخزن قاعدة البيانات مفاتيح الكائنات لا روابط عامة دائمة. ينشئ الخادم روابط مؤقتة بعد التحقق من الصلاحية عند الحاجة إلى التنزيل. ويطبق المبدأ نفسه على ملف المقال قبل النشر، بينما تصبح إتاحة المقال المنشور عملية مضبوطة من خلال واجهات النشر العامة.

تستخدم المصادقة رموز وصول، إلا أن وجود الرمز لا يكفي وحده؛ فكل عملية تحريرية تعيد فحص الدور والنطاق والحالة. كما لا تسجل رسائل فشل كشف الانتحال مسارات محلية أو بيانات اتصال أو آثاراً تفصيلية يمكن أن تكشف بيئة التشغيل، بل تحول الأخطاء إلى رموز ورسائل آمنة للمستخدم.

6.8 تصميم المعالجة غير المتزامنة والموثوقية

تنفذ الطلبات القصيرة والتفاعلية داخل خادم التطبيق، بينما تحول الأعمال الطويلة أو الدورية إلى طابور مهام. وتشمل الأعمال الخلفية إرسال الإشعارات، وانتهاء دعوات المحكّمين، وتذكيرات المواعيد، وتحديث بعض بيانات المحكّمين، والتحقق من الحاجة إلى نمذجة موضوعات القسم، وفحص الانتحال.

يفصل التصميم بين الطابور العام وطابور الانتحال. يعمل العامل العام بقدرة تزامن تناسب المهام القصيرة، بينما يعمل عامل الانتحال بتزامن منخفض وجلب مسبق محدود، لأن تحميل نماذج E5 وAraT5 والفهارس يستهلك ذاكرة أكبر. ويعاد استخدام خط الانتحال داخل عملية العامل بدلاً من تحميله لكل مخطوطة.

تسجل العملية التحريرية الأساسية أولاً، ثم ترسل آثارها الجانبية بعد نجاح المعاملة باستخدام إجراء ينفذ بعد التثبيت. فإذا سجل قرار المحرر بنجاح وفشل إرسال البريد، لا يتراجع القرار ولا تتناقض قاعدة البيانات مع رسالة أرسلت قبل اكتمال المعاملة. وبالمثل لا يبدأ فحص الانتحال قبل ثبات إصدار المخطوطة وملفه.

تستخدم الخدمات التي تعدل عدة سجلات معاملات ذرية وأقفال صفوف عند وجود تنافس محتمل. من الأمثلة تسلسل استجابات المحكّمين للإصدار نفسه حتى لا يتجاوز عدد المقبولين الحد المطلوب، وقفل المخطوطة عند اكتمال التقارير والانتقال إلى حالة المراجعة. ولا يفترض التصميم أن كل مهمة في النظام مكررة التنفيذ بأمان بالدرجة نفسها؛ بل تطبق الحماية الأقوى حيث يوجد خطر مثبت، كما في سجل فحص الانتحال وحالاته المقيدة.

6.9 تصميم إعادة الاستخدام والتخصيص

صمم النظام لتشغيل مجلة واحدة في كل عملية تنصيب، مع تهيئته لمؤسسة أو مجلة مختلفة من دون إعادة بناء النواة التحريرية. يمثل JournalMetadataSettings إعداداً وحيداً على مستوى التنصيب، ويحتفظ باسم المجلة واختصارها وشعارها ولونها ووصفها والناشر وأرقام ISSN واللغة والرخصة وسياسة الوصول والتحكيم وبعض بيانات OAI وأوزان ترتيب الأولوية.

يمكن للجهة إنشاء أقسامها وتعيين مديريها وعضويات محرريها، وإدارة صفحات عامة محددة مثل التعريف بالمجلة وإرشادات المؤلفين وأخلاقيات النشر والوصول المفتوح وهيئة التحرير والتواصل. اختيرت أنواع صفحات ثابتة بدلاً من بناء نظام إدارة محتوى عام؛ فهي تغطي نطاق المجلة وتحافظ على بساطة التحقق والعرض.

لا تعد جميع عناصر النظام قابلة للتخصيص. فقائمة الأدوار الجوهرية ودلالات حالات المخطوطة وعلاقة التقرير بالإصدار وقواعد حفظ التاريخ تبقى جزءاً من النواة، لأن تغييرها عشوائياً قد يكسر سير العمل أو الصلاحيات. أما الأشخاص المنتمون إلى الأدوار، ونطاقاتهم في الأقسام، وهوية المجلة ومحتواها وأوزان الأولوية وبيانات النشر، فهي قابلة للتهيئة.

قابل للتهيئة	ثابت في النواة
اسم المجلة وشعارها وألوانها وبياناتها الوصفية	الفصل بين المخطوطة وإصداراتها
الأقسام ومديروها وعضويات المحررين	انتقالات الحالات القانونية
صفحات السياسات والمحتوى العام	ربط التحكيم بالإصدار الذي تمت مراجعته
بيانات الأعداد والنشر والرخصة	الفصل بين المخطوطة المقبولة والمقال المنشور
أوزان عوامل ترتيب الأولوية	بقاء مخرجات الذكاء الاصطناعي دعماً للقرار

يطبق ذلك مبدأ **الإعدادات بدلاً من الشيفرة (Configuration over Code)** بصورة مقيدة: تنقل الفروق المؤسسية المناسبة إلى البيانات والإعدادات، مع إبقاء قواعد سلامة المجال في الشيفرة المشتركة.

6.10 الأنماط والأساليب التصميمية المعتمدة

يستخدم النظام أنماطاً على مستويات مختلفة، ولذلك من المهم عدم جمعها كلها تحت مسمى أنماط GoF. يبين الجدول الآتي التصنيف الأدق وموضع التطبيق:

العنصر	التصنيف	موضعه في النظام
طبقة الخدمات Service Layer	نمط تطبيقي	عمليات الفرز والإسناد والتحكيم والقرار والنشر والتحليلات
خط المعالجة Processing Pipeline	نمط معالجة	المراحل المتتابعة لكشف الانتحال
المستودع Repository	نمط وصول إلى البيانات	PostgresSourceRepository لمصادر ومقاطع الانتحال
الاستراتيجية Strategy	نمط سلوكي	عقود المعالج والمسترجع والمتحقق المحقونة في منسق الانتحال
المكيّف Adapter	نمط بنيوي	ProductionCandidateAdapter بين ناتج الاسترجاع ومدخل المتحقق
دالة مصنع/جذر تجميع	أسلوب إنشاء وتهيئة	build_default_local_pipeline لتجميع مكونات خط الانتحال وفق الإعدادات
المنتج والمستهلك	نمط تزامن وتكامل	إنتاج مهام Celery واستهلاكها من العوامل
آلة حالات صريحة	أسلوب ضبط لسير العمل	جدول انتقالات المخطوطة والخدمات التي تطبقه

تفصل **طبقة الخدمات** قواعد العمليات المركبة عن واجهات API. فالواجهة تستقبل الطلب وتتحقق من بنيته، ثم تستدعي خدمة تنفذ القواعد والمعاملة والتغييرات المترابطة. يمنع ذلك تكرار القواعد بين عدة مسارات، ويجعل الانتقال من حالة إلى أخرى نتيجة لعملية مجال كاملة لا لتعديل حقل مباشر.

يظهر **نمط المستودع** بوضوح داخل منظومة الانتحال فقط، ولا يدعي التصميم وجود مستودع عام فوق كل نماذج Django. يعزل PostgresSourceRepository استعلامات مصادر المقارنة عن الاسترجاع والفهرسة. وبالمثل تسمح عقود المعالجة والاسترجاع والتحقق باستبدال تنفيذ إحدى المراحل من دون تعديل منسق الخط، وهو تطبيق لنمط **الاستراتيجية**. أما المكيّف فيحل اختلافاً فعلياً بين بنية المرشحين وبنية مدخلات المتحقق.

تعد دالة بناء خط الانتحال نقطة مركزية لتجميع المستودع والمعالج والمسترجعين والمتحقق وباني التقرير. وهي أقرب إلى **دالة مصنع بسيطة وجذر تجميع (Composition Root)** منها إلى تطبيق كلاسيكي لنمط Factory Method. أما انتقالات المخطوطة فهي آلة حالات صريحة وجدول انتقالات مركزي، وليست تطبيقاً لنمط State ذي كائن مستقل لكل حالة.

6.11 تصميم النشر والتوزيع

تنشر مكونات الخادم في حاويات مترابطة ضمن شبكة تشغيل واحدة. توجد حاوية لقاعدة PostgreSQL المدعومة بـpgvector، وحاوية مستقلة لتنفيذ الترحيلات قبل بدء التطبيق، وحاوية للخادم الخلفي، وRedis وسيطاً للمهام، وعامل Celery عام، وعامل مخصص لكشف الانتحال، ومجدول Celery Beat، وMinIO لتخزين الملفات.

تفصل البيانات الدائمة عن دورة حياة الحاويات باستخدام وحدات تخزين لقاعدة البيانات وRedis وMinIO وذاكرة BM25. وتثبت موارد نماذج الانتحال والفهرس الدلالي في العامل المخصص للقراءة فقط، بينما تخصص له مساحة قابلة للكتابة للذاكرة المعجمية. يقلل هذا الفصل خطر تعديل النماذج أثناء التشغيل، ويمنع فقد البيانات عند إعادة إنشاء الحاوية.

تبقى قاعدة البيانات وRedis ومخزن الملفات مكونات داخلية، ويصل المستخدم إليها عبر الخادم فقط. وتتصل خدمات البريد وORCID خارج حدود النظام وفق الحاجة، بينما يتيح OAI-PMH للأنظمة الحاصدة قراءة البيانات الوصفية العامة المنشورة لا البيانات التحريرية.

تمثل واجهة Next.js مكوناً مستقلاً في التصميم المنطقي، إلا أن ملف Docker Compose الحالي يركز على الخدمات الخلفية ويترك تعريف حاوية الواجهة غير مفعل. لذلك يجب ألا يصور مخطط النشر الحالي الواجهة على أنها حاوية نشطة ما لم يوضح هذا الفرق. كما لا يتضمن التصميم الحالي موازن أحمال أو Kubernetes أو Nginx، فلا تضاف هذه العناصر إلى المخطط بوصفها منفذة.

> **ملاحظة لإدراج شكل (6-8):** يوضع هنا **مخطط نشر (Deployment Diagram)** يعرض المتصفح، وواجهة Next.js كمكوّن مستقل مع ملاحظة أنها غير مفعلة في Compose الحالي، وحاويات backend وmigrate وPostgreSQL/pgvector وRedis وcelery_worker وplagiarism_worker وcelery_beat وMinIO، إضافة إلى وحدات التخزين الدائمة وخدمة البريد الخارجية. لا تضاف مكونات بنية تحتية غير موجودة في المشروع.

الفصل السابع

الأدوات المستخدمة

يعرض هذا الفصل الأدوات المستخدمة لتنفيذ المشروع

1.7 Django و Django REST Framework

يعد Django إطار عمل لتطوير تطبيقات الويب بلغة Python، بينما يوفر Django REST Framework (DRF) أدوات لبناء واجهات REST البرمجية. استخدما في المشروع لتطوير الواجهة الخلفية وإدارة منطق الأعمال والبيانات والمستخدمين والصلاحيات، إضافة إلى توفير واجهات برمجية تربط الواجهة الأمامية بخدمات النظام. كما استخدمت JWT للمصادقة و-drf spectacular لتوليد توثيق OpenAPI للواجهات البرمجية.

2.7 PostgreSQL و pgvector

تعد PostgreSQL قاعدة بيانات علائقية مفتوحة المصدر، بينما تضيف pgvector دعماً لتخزين المتجهات وإجراء عمليات التشابه عليها داخل قاعدة البيانات. استخدمت PostgreSQL لتخزين بيانات النظام، مثل المستخدمين والمخطوطات وعمليات التحكيم والنشر، في حين استخدمت pgvector لتخزين التمثيلات الدلالية والاستفادة منها في عمليات المطابقة، وخصوصاً في توصية المحكّمين.

3.7 Celery و Redis

يوفر Celery آلية لتنفيذ المهام غير المتزامنة وخارج دورة طلب المستخدم، بينما يستخدم Redis كوسيط رسائل لنقل المهام بين التطبيق وعمال التنفيذ. استخدما في المشروع لتنفيذ العمليات طويلة أو دورية التنفيذ، مثل توليد التمثيلات الدلالية، وتجميع المخطوطات موضوعياً، وإرسال التذكيرات، وتشغيل فحص الانتحال. كما استخدم Celery Beat لجدولة بعض هذه المهام بصورة دورية.

4.7 MinIO

يعد MinIO نظاماً لتخزين الكائنات متوافقاً مع واجهات Amazon S3، ويستخدم لتخزين الملفات بصورة منفصلة عن قاعدة البيانات. استخدم في المشروع لتخزين ملفات المخطوطات وإصداراتها المختلفة، مع توفير روابط موقعة مؤقتة للوصول إليها، إضافة إلى تنزيلها مؤقتاً عند الحاجة إلى إجراء عمليات معالجة مثل استخراج النص أو فحص الانتحال.

5.7 React و Next.js

تعد React مكتبة لبناء واجهات المستخدم التفاعلية بالاعتماد على مكونات قابلة لإعادة الاستخدام، بينما يعد Next.js إطار عمل مبنياً عليها يوفر تنظيم الصفحات والمسارات وآليات بناء تطبيقات الويب. استخدم React 19 وNext.js 16 لتطوير الواجهة الأمامية للنظام وبناء الصفحات العامة وبوابات المستخدمين المختلفة بحسب أدوارهم.

6.7 Axios و TanStack Query

تستخدم TanStack Queryلإدارة البيانات القادمة من الخادم وحالات التحميل والتخزين المؤقت وتحديث البيانات، بينما توفر Axiosواجهة مبسطة لتنفيذ طلبات HTTP. استخدمتا في المشروع للتواصل مع واجهات Django REST API وإدارة البيانات المعروضة في الواجهة الأمامية وتحديثها بعد تنفيذ العمليات المختلفة.

7.7 Tailwind CSS و shadcn/ui

يعد Tailwind CSSإطاراً لتنسيق واجهات الويب باستخدام أصناف CSS جاهزة وقابلة للتركيب، بينما توفر shadcn/ui مجموعة من مكونات الواجهة القابلة للتخصيص. استخدما لبناء واجهات النظام وتوحيد مظهر عناصرها مع الحفاظ على إمكانية إعادة استخدام المكونات وتخصيصها.

8.7 BERTopic و UMAP

تعد BERTopicمكتبة لاكتشاف الموضوعات وتجميع الوثائق النصية اعتماداً على تمثيلاتها الدلالية، بينما تستخدم UMAP لاختزال أبعاد هذه التمثيلات قبل عملية التجميع. استخدمتا في المشروع لتجميع المخطوطات داخل الأقسام بحسب موضوعاتها واستخراج الكلمات الممثلة لكل موضوع للاستفادة منها في التحليلات ودعم نظام التوصية.

9.7 PyTorch و Hugging Face Transformers

يعد PyTorch إطاراً لبناء وتدريب وتشغيل نماذج التعلم العميق، بينما توفر Hugging Face Transformers أدوات ونماذج جاهزة مبنية على معماريات المحولات. استخدما بصورة رئيسة في خط كشف الانتحال لتشغيل نموذج multilingual-E5 للاسترجاع الدلالي، وتدريب وتشغيل نموذج AraT5v2 للتحقق من أزواج المقاطع المرشحة.

10.7 Docker

يوفر Docker آلية لتشغيل التطبيقات وخدماتها ضمن حاويات معزولة، بينما يسمح Docker Compose بإدارة وتشغيل مجموعة من الحاويات المترابطة من خلال إعداد موحد. استخدما في المشروع لتشغيل خدمات الواجهة الخلفية و PostgreSQL و Redis و MinIO وعمال Celery وخدمات فحص الانتحال ضمن بيئة موحدة، مما سهّل إعداد النظام وتشغيله بصورة متكررة على أجهزة مختلفة.

11.7 ORCID

تعد ORCID خدمة توفر معرفات رقمية دائمة للباحثين وتربطها بمعلوماتهم ومنشوراتهم العلمية. دمجت الخدمة في المشروع لاسترجاع بيانات منشورات المحكّمين عند توفر معرف ORCID، واستخدام عناوين هذه المنشورات مع الكلمات المفتاحية والسيرة العلمية لبناء تمثيل أكثر شمولاً لخبرة المحكّم والاستفادة منه في نظام التوصية

12.7 Grafana k6

تعد Grafana k6 أداة مفتوحة المصدر لاختبار أداء تطبيقات الويب وواجهات API من خلال محاكاة عدد من المستخدمين والطلبات المتزامنة. استخدمت في المشروع لإجراء اختبارات الحمل والأداء على واجهات REST، وقياس مؤشرات مثل زمن الاستجابة، ومعدل الطلبات، ونسبة الأخطاء، ومراقبة سلوك النظام عند زيادة الحمل للتحقق من قدرته على التعامل مع الاستخدام المتزامن.

13.7 Git

يعد Git نظاماً موزعاً للتحكم بإصدارات الشيفرة المصدرية وتتبع التعديلات عليها. استخدم أثناء تطوير المشروع لحفظ تاريخ التغييرات، وإنشاء فروع مستقلة لتطوير الميزات المختلفة، ثم دمجها ضمن النسخة الرئيسة بعد الانتهاء منها.

الفصل الثامن

تنفيذ النظام

يعرض هذا الفصل كيفية تنجيز النظام مع تفاصيل اجزائه

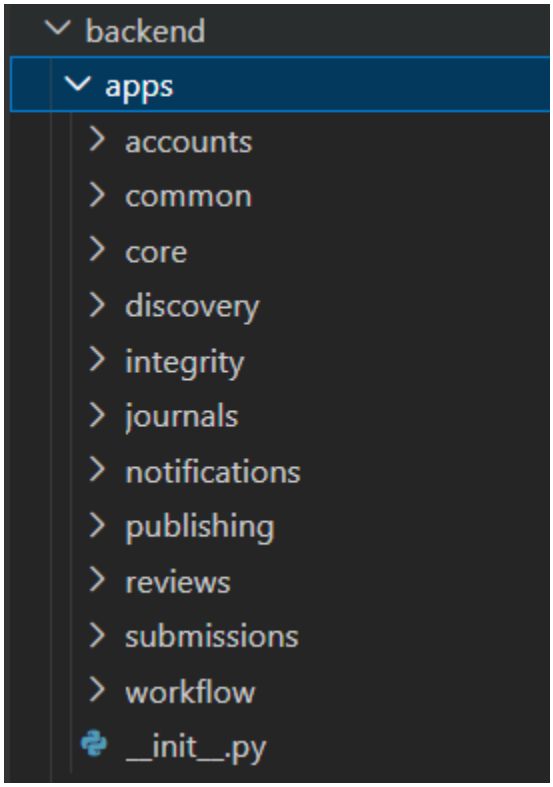
1.8 مقدمة

يعرض هذا الفصل كيفية تحويل التصميم المعتمد للنظام الذكي لإدارة المجلات العلمية المحكمة (Intelligent Journal Management System – IJMS) إلى مكونات برمجية مترابطة تدعم دورة النشر العلمي من إرسال المخطوطة حتى نشر المقالة وإتاحتها للقراء. ويركز العرض على الوحدات التي نُفذت فعلياً، وطريقة تفاعلها مع البيانات والمهام غير المتزامنة وواجهات المستخدم، من دون إعادة مناقشة المبادئ النظرية أو مبررات اختيار المعمارية والأدوات، ومن دون الدخول في نتائج الاختبارات التي ستعرض في الفصل التالي.

نُفذ النظام على شكل تطبيق ويب ذي واجهة خلفية مبنية باستخدام Django و Django REST Framework، وواجهة أمامية مبنية باستخدام Next.js. وتولت PostgreSQL حفظ البيانات المنظمة والمتجهات، بينما استُخدم MinIO لحفظ ملفات المخطوطات، وتولى Celery تنفيذ الأعمال الخلفية بالاعتماد على Redis. وفوق الوظائف التحريرية الأساسية، دُمجت مكونات توصية المحكمين ونمذجة الموضوعات وترتيب الأولويات والتحليلات وكشف الانتحال بوصفها أدوات مساندة للقرار التحريري، لا بدائل عن الحكم البشري.

2.8 تنظيم المشروع وتنفيذ البنية الخلفية

نُظمت الواجهة الخلفية في تطبيقات نطاقية مستقلة داخل مشروع Django واحد. تضم قائمة التطبيقات المحلية الوحدات الموضحة بالشكل X، تتولى كل وحدة جزءاً محدداً من دورة العمل، مع بقاء العلاقات الضرورية بينها واضحة؛ فمثلاً تحفظ وحدة `submissions` المخطوطة وإصداراتها، بينما تنفذ وحدة `workflow` التعيين والفرز الأولي وانتقالات الحالة، وتدير وحدة `reviews` الدعوات وتقارير التحكيم، ثم تنشئ وحدة `publishing` سجل النشر اعتماداً على الإصدار المقبول.



نُفذت واجهة البرمجة باستخدام Django REST Framework. استُخدمت المسلسلات (Serializers) للتحقق من بيانات الطلبات وصياغة الاستجابات، وطبقات الصلاحيات لتقييد الوصول، في حين وُضعت العمليات التي تجمع عدة قواعد عمل في خدمات نطاقية مستقلة عن العروض (Views)، مثل `SubmissionService` و`AssignmentService` و`ReviewService` و`PublishingService`. أتاح ذلك إبقاء العروض مسؤولة عن التعامل مع HTTP، مع تنفيذ القواعد التحريرية وتحديث البيانات داخل خدمات يمكن استدعاؤها من أكثر من واجهة.

تعتمد واجهات REST على مصادقة الرموز المميزة من نوع JWT، بينما تُطبق الصلاحيات على مستويين: صلاحية عامة مرتبطة بدور المستخدم، وصلاحية سياقية مرتبطة بالمجلة أو القسم أو المخطوطة المعنية. كما تستخدم الخدمات معاملات قاعدة البيانات وقفل الصفوف في العمليات الحساسة، وتؤجل الآثار الجانبية، مثل إرسال البريد أو تشغيل مهمة خلفية، إلى ما بعد نجاح تثبيت المعاملة. أما الملفات فلا تحفظ داخل قاعدة البيانات، بل تُرفع إلى حاويات خاصة في MinIO، وتبقى قاعدة البيانات محتفظة بمفاتيح الكائنات والبيانات الوصفية اللازمة للوصول المضبوط إليها.

[شكل مقترح: بنية التطبيقات الرئيسية في الواجهة الخلفية، مع بيان التطبيقات المسؤولة عن الحسابات والمخطوطات وسير العمل والتحكيم والنشر والاكتشاف والنزاهة العلمية]

8.3 تنفيذ إدارة المستخدمين والأدوار

يمثل نموذج المستخدم نقطة الارتباط بين الهوية الرقمية والأدوار التحريرية. نُفذت الأدوار على شكل علاقة متعددة القيم، بحيث يستطيع الحساب الواحد أن يكون مؤلفاً ومحكّماً أو أن يحمل دوراً تحريرياً إضافياً من دون إنشاء حسابات منفصلة. وتشمل الأدوار المنفذة: المؤلف، والمحكّم، ومدير القسم، ومحرر القسم، ورئيس التحرير، ومدير النظام، إضافة إلى القارئ. ويحتفظ الملف الشخصي ببيانات مثل الاسم والبريد والانتماء المؤسسي والدولة ومعرّف ORCID عند توفره.

لأن بعض الصلاحيات لا يكفي ضبطها بدور عام، أضيفت عضوية محرر القسم لربط المحرر بالأقسام التي يعمل ضمنها. وبذلك لا يتمكن محرر القسم من إدارة كل مخطوطات المجلة، بل المخطوطات المسندة إليه وفي النطاق المسموح له. وبالمثل يرتبط مدير القسم بقسم محدد، وتتحقق الخدمات من هذا الارتباط عند الفحص الأولي أو متابعة التحكيم أو النشر. أما رئيس التحرير ومدير النظام فيحصلان على نطاق أوسع يلائم إدارة المجلة بأكملها.

نُفذ الانضمام إلى هيئة التحكيم كعملية اعتماد مستقلة. يرسل المستخدم طلب محكّم لقسم واحد مع كلمات مفتاحية وسيرة مختصرة، ثم يراجعه مدير القسم أو الجهة المخولة ويقبله أو يرفضه مع تسجيل المراجع والملاحظة. وعند القبول يُنشأ أو يُحدّث ملف المحكّم، ويُربط بالقسم، وتُحفظ كلمات خبرته وبيانات مزامنة ORCID وتمثيل الخبرة الدلالي. ويفصل هذا الإجراء بين امتلاك حساب في النظام وبين الأهلية الفعلية لاستلام دعوات التحكيم.

8.4 تنفيذ إرسال المخطوطات وإدارة الإصدارات

تبدأ عملية الإرسال ببيانات المخطوطة الأساسية، وتشمل العنوان والملخص والكلمات المفتاحية واللغة والقسم وخطاب التغطية، إلى جانب بيانات المؤلفين المشاركين وترتيبهم. يرفع المؤلف نسختين: نسخة كاملة تحفظ معلومات التأليف، ونسخة مجهّلة مخصصة للتحكيم. تُخزن النسختان في MinIO تحت مفاتيح لا تُعرض مباشرة للمستخدم، ويُنشأ في قاعدة البيانات سجل مخطوطة مرتبط بالإصدار الأول.

نُفذت خدمة الإرسال بصورة ذرّية قدر الإمكان. فهي تتحقق من المدخلات، وتنشئ سجلات المخطوطة والمؤلفين والإصدار ضمن معاملة واحدة، وتزيل الملفات المرفوعة على نحو أفضل جهد ممكن إذا تعذر إكمال إنشاء السجلات. وبعد تثبيت المعاملة، تُرسل أعمال الاشتقاق التي لا ينبغي أن تؤخر استجابة المستخدم، مثل تكوين تمثيل الملخص الدلالي، كما يُطلب الفحص الأولي للتشابه تلقائياً للإصدار العربي الأول عندما تكون منظومة الفحص مفعّلة ومهيأة.

لا تُستبدل ملفات المخطوطة عند التعديل، بل يُنشأ SubmissionVersion جديد يحمل رقماً متسلسلاً وملفيه الكامل والمجهّل ورسالة الرد على المحكّمين. ويحتفظ كل إصدار بقراره وخطاب القرار ومن اتخذه وتوقيته، مما يجعل تاريخ المخطوطة قابلاً للتتبع. ولا يسمح برفع مراجعة جديدة إلا لصاحب المخطوطة، وفي حالة طلب تعديل فعلي، وضمن الحد الأقصى لجولات المراجعة المعرّف في إعدادات المجلة.

عند رفع إصدار منقح، تعاد المخطوطة إلى مرحلة التحكيم، وتُنشأ دورة مرتبطة بالإصدار الجديد. وإذا كان محكّم من الدورة السابقة قد أكمل تقريره وقُبل للاستمرار، تُرحّل علاقته إلى الإصدار الجديد مع الاحتفاظ بمرجع إلى التكليف السابق وتحديد موعد جديد. يحقق ذلك الاستمرارية بين الجولات من دون خلط تقارير الإصدارات أو فقدان تاريخها.

8.5 تنفيذ سير العمل التحريري

نُفذت انتقالات المخطوطة في وحدة مركزية تحدد الحالات المسموح الانتقال بينها. تبدأ المخطوطة بحالة «مرسلة»، ثم يمكن أن تُسند أو تُرفض مكتبياً. وبعد الإسناد تنتقل إلى «قيد التحكيم»، ثم إلى «تمت مراجعتها» عند اكتمال التقارير، وبعدها يصدر قرار قبول أو رفض أو طلب تعديل. وتعود المخطوطة المعدلة إلى التحكيم، في حين تعد حالتا القبول والرفض نهائيتين في السجل التحريري. يمنع هذا الضبط تغيير الحالة إلى قيمة صحيحة لغوياً لكنها غير منطقية بالنسبة إلى موقع المخطوطة في الدورة.

يبدأ العمل التحريري الفعلي بالفحص الأولي الذي ينفذه مدير القسم المسؤول. تتضمن قائمة الفحص بنوداً إلزامية تغطي ملاءمة النطاق، واكتمال المخطوطة والبيانات الوصفية، والالتزام بإرشادات المؤلفين، والجودة الأساسية، والإفصاحات الأخلاقية. يستطيع المدير حفظ القائمة كمسودة، أما إكمال الفحص بالموافقة فيتطلب عدم وجود ملاحظات مانعة. وفي حالة الرفض المكتبي يجب تحديد سبب وملاحظة للمؤلف، ويسجل النظام النتيجة ومن اتخذها ووقت اتخاذها.

بعد اجتياز الفحص، يسند مدير القسم المخطوطة إلى محرر قسم مناسب. يحتفظ سجل الإسناد بالمستخدم الذي نفذ العملية، والمحرر المسند إليه، وصورة الدور وقت الإسناد، والسبب، كما يسمح بإعادة الإسناد مع بقاء التاريخ السابق. ويعتمد ما يلي من دعوات وقرارات على المحرر المسند فعلياً، لا على مجرد امتلاك المستخدم لدور محرر قسم.

تُنفذ كل عملية انتقال من خلال خدمة تتحقق أولاً من الحالة والصلاحية والبيانات التابعة، ثم تحدث السجلات ذات الصلة داخل معاملة واحدة. فعلى سبيل المثال، لا يقتصر قرار التعديل على تغيير حالة المخطوطة، بل يرتبط بقرار الإصدار وخطابه والمستخدم الذي اتخذه. وبالمثل يؤدي اكتمال جميع تقارير التحكيم المطلوبة إلى تحديث حالة المخطوطة بطريقة منضبطة، بدلاً من ترك الواجهة الأمامية تستنتج الحالة أو تعدلها مباشرة.

> [مخطط مقترح: التدفق التنفيذي للمخطوطة من الإرسال والفحص الأولي إلى الإسناد والتحكيم والقرار، مع إظهار نقاط التحقق من الصلاحيات والحالة]

8.6 تنفيذ نظام التحكيم العلمي

يمثل ReviewerAssignment دعوة محكّم لإصدار محدد، ويحتفظ بحالتها وموعد الاستجابة وموعد تسليم التقرير، إضافة إلى بيانات الإلغاء أو الاستبدال والترحيل من إصدار سابق. وتنتقل الدعوة بين حالات الانتظار والقبول والاعتذار والانتهاء والإلغاء. أما تقرير التحكيم فيرتبط بالتكليف المقبول، ويفصل بين الملاحظات الموجهة للمؤلف والملاحظات السرية للمحرر، ويخزن التوصية التحريرية ضمن الخيارات المعتمدة.

يستطيع محرر القسم المسند دعوة محكّمين معتمدين في قسم المخطوطة. تتحقق الخدمة من أن المرشح ليس المؤلف أو أحد المؤلفين المشاركين، وأنه لم يُدع في الدورة نفسها، وأن عبءه الحالي لم يتجاوز الحد، وأن ملفه معتمد للقسم. ويمكن إرسال عدة دعوات ضمن عملية واحدة، ثم يستجيب كل محكّم من مساحة عمله. وعند قبول دعوة، يُقفل سجل الإصدار أثناء التحقق من عدد المقبولين حتى لا تؤدي الاستجابات المتزامنة إلى تجاوز العدد المطلوب؛ وإذا اكتمل العدد تُنهى الدعوات المعلقة الزائدة.

يعرض النظام للمحكّم النسخة المجهّلة فقط بعد قبول التكليف، ولا يكشف له بيانات المؤلفين. وفي الاتجاه المقابل لا يعرض للمؤلف هوية المحكّم، كما لا يدمج الملاحظات السرية ضمن خطاب القرار. وتتحقق واجهات تنزيل الملفات من علاقة المستخدم بالمخطوطة ونوع الملف المطلوب بدلاً من الاعتماد على إخفاء الروابط وحده، وهو ما يطبق التحكيم المزدوج المجهول في طبقة البيانات والخدمة والواجهة معاً.

بعد إرسال التقرير، يراقب النظام اكتمال تقارير التكليفات المقبولة. وعندما تكتمل جميعها تصبح المخطوطة جاهزة للقرار، ويستطيع محرر القسم المسند إصدار القبول أو الرفض أو التعديل البسيط أو الجوهري مع خطاب قرار. لا يسمح بإصدار القرار قبل اكتمال التقارير، ولا يسمح بتجاوز الحد الأقصى لجولات التعديل. كما تنفذ مهام دورية لإنهاء الدعوات التي تجاوزت مهلة الاستجابة وإرسال تذكيرات بالمواعيد النهائية للتقارير.

8.7 تنفيذ النشر والأرشفة والاكتشاف

فُصل السجل المنشور عن سجل المخطوطة التحريري. بعد قبول المخطوطة يمكن إنشاء مسودة PublishedArticle انطلاقاً من الإصدار المقبول، فتُنسخ إليها بيانات العنوان والملخص والكلمات المفتاحية واللغة والمؤلفين وترتيبهم وملف الإصدار المصدر. يسمح هذا النسخ بتثبيت البيانات العامة للمقالة وعدم تأثرها لاحقاً بتغير ملف مؤلف أو بيانات تشغيلية في المخطوطة، مع إبقاء مرجع واضح إلى الإصدار الذي اعتمد للنشر.

تظل المسودة غير ظاهرة للعامة إلى أن تُستكمل بيانات النشر، مثل المعرّف الرقمي عند توفره، والترخيص، والصفحات، والعدد. يسمح النظام بإنشاء المسودة للمستخدمين التحريريين المرتبطين بالمخطوطة، لكنه يقيد تنفيذ النشر النهائي برئيس التحرير أو مدير النظام أو مدير القسم المسؤول. ويشترط أن يوجد عدد جارٍ مفتوح للنشر. وعند النشر تسجل لحظة الإتاحة، وترتبط المقالة بالعدد والقسم، ثم تُرسل رسالة إلى المؤلف المراسل.

نُفذت إدارة الأعداد بحالات «مسودة» و«منشور جارٍ» و«مؤرشف»، مع قيد يمنع وجود أكثر من عدد جارٍ في الوقت نفسه. يتطلب فتح العدد تحديد العنوان والمجلد والرقم، ويستقبل المقالات المنشورة تباعاً. أما إغلاقه فينقله إلى الأرشيف، ولا يسمح بذلك إذا بقيت فيه مسودات مقالات غير منشورة، مما يحافظ على اتساق محتوى العدد.

توفر الواجهة العامة قوائم المقالات المنشورة وصفحات تفاصيلها والبحث في بياناتها، إضافة إلى التصفح حسب الأقسام والأعداد والأرشيف. ولا تُعرض المسودات أو المقالات غير المنشورة في هذه الواجهات. يبقى ملف المقالة داخل مخزن خاص، ويصدر الخادم عند طلب التنزيل رابطاً موقعاً محدود الصلاحية بعد التحقق من حالة المقالة، بدلاً من جعل حاوية الملفات عامة.

دُعمت قابلية الاكتشاف من خلال تصدير البيانات الوصفية بصيغ BibTeX وRIS وDublin Core، ومن خلال واجهة OAI-PMH. تنفذ الواجهة أوامر Identify وListMetadataFormats وListSets وListIdentifiers وListRecords وGetRecord، وتعرض السجلات المنشورة فقط بصيغة oai_dc. كما أضيفت صفحات ثابتة مضبوطة الأنواع لعرض معلومات المجلة، مثل نبذة المجلة وإرشادات المؤلفين وأخلاقيات النشر وسياسة الوصول المفتوح وهيئة التحرير وصفحة التواصل، ولا تظهر للعامة إلا بعد نشر محتواها.

> [صورة مقترحة: مساحة رئيس التحرير لإدارة العدد الجاري ومسودات المقالات وحالة جاهزية كل مقالة للنشر]

8.8 تنفيذ نظام توصية المحكّمين

دُمجت التوصية في مساحة عمل محرر القسم عند اختيار المحكّمين، ونُفذت في RecommendationService بوصفها خدمة تقرأ المخطوطة وملفات المحكّمين وتعيد قائمة مرشحين مرتبة مع تفسير مختصر. تبدأ الخدمة بتصفية المرشحين وفق قواعد الأهلية التنفيذية: اعتماد المحكّم في قسم المخطوطة، وامتلاكه الدور المناسب، وعدم كونه مؤلفاً أو مؤلفاً مشاركاً، وعدم وجود دعوة أو تكليف نشط له للمخطوطة نفسها، وعدم بلوغ الحد الأقصى لعبء التحكيم.

بعد التصفية تحسب الخدمة دليلين للملاءمة. الدليل الأول هو التشابه الدلالي بين تمثيل ملخص المخطوطة وتمثيل خبرة المحكّم المخزنين في حقول متجهية في PostgreSQL، ويُرتب باستخدام مسافة جيب التمام التي يوفرها pgvector. أما الدليل الثاني فيقيس تغطية كلمات المخطوطة بكلمات خبرة المحكّم، وتدخل فيه الكلمات التي أدخلها المؤلف والكلمات الناتجة عن التحليل الموضوعي عند توفرها. ثم تُجمع القيمتان بأوزان قابلة للضبط في الإعدادات؛ وإذا غاب أحد الدليلين يعاد استخدام الوزن المتاح بدلاً من إسناد قيمة مصطنعة للدليل المفقود.

لا تُرجع الخدمة درجة مجردة فقط، بل تبين نمط الاحتساب والكلمات المتطابقة والعبء الحالي وما إذا كان للمحكّم خبرة تحكيم سابقة. ويجري ترتيب النتائج بالدرجة أولاً، ثم بالعبء الأقل والاسم عند التعادل. يختار المحرر من القائمة أو يتجاوزها، وتبقى عملية إرسال الدعوة خاضعة للتحققات نفسها المستخدمة مع الاختيار اليدوي؛ ولذلك لا تعني التوصية تعيين المحكّم تلقائياً.

> [صورة مقترحة: واجهة محرر القسم أثناء استعراض المحكّمين المقترحين، مع الدرجة والكلمات المتطابقة وعبء التحكيم وتفسير الترتيب]

8.9 تنفيذ نمذجة الموضوعات واستخراج الكلمات المفتاحية

نُفذ التحليل الموضوعي كمهمة دورية غير متزامنة على مستوى القسم. تفحص مهمة مجدولة الأقسام الفعالة، وتطلق إعادة العنقدة عندما تتوافر مخطوطات جديدة ذات ملخصات وتمثيلات دلالية. ويمنع التنفيذ الحالي تشغيل التحليل على مجموعة أصغر من الحد البرمجي الأدنى، وهو ثلاث مخطوطات، لأن التحليل على حالة منفردة أو عدد غير كافٍ لا ينتج تجميعاً ذا معنى.

تستخدم المهمة التمثيلات الدلالية المحسوبة مسبقاً للملخصات بدلاً من إعادة توليدها في كل تشغيل. تُمرر هذه التمثيلات إلى UMAP للاختزال، ثم إلى BERTopic لتكوين الموضوعات. وبعد انتهاء العنقدة تستخرج الخدمة الكلمات الأعلى تمثيلاً لكل موضوع من ناتج BERTopic، وتحفظ لكل مخطوطة تسمية الموضوع وقائمة الكلمات في SubmissionTopic. أما المخطوطات التي يصنفها النموذج حالات شاذة فتُحفظ لها قيمة موضوع فارغة وقائمة كلمات فارغة، كي لا تُعرض نتيجة غير مؤكدة على أنها موضوع حقيقي.

ترتبط المخرجات ببقية النظام بطريقتين: تُضاف الكلمات الموضوعية إلى الأدلة المستخدمة عند توصية المحكّمين، وتُجمع الموضوعات في لوحة التحليلات لعرض توزيع محتوى المجلة. كما يسجل القسم وقت آخر عنقدة، مما يتيح للمهمة المجدولة تحديد الحاجة إلى إعادة التشغيل من دون إجراء العمل الثقيل بعد كل طلب ويب.

8.10 تنفيذ دعم القرار والتحليلات التحريرية

8.10.1 ترتيب المخطوطات حسب الأولوية

نُفذت خدمة SubmissionPriorityService لتكوين قائمة عمل مرتبة للمخطوطات التي تحتاج إلى متابعة، وليس للتنبؤ بجودتها أو احتمال قبولها. تقرأ الخدمة الحالة الحالية وسجل الأزمنة والتكليفات وجولات المراجعة، ثم تحسب درجة من صفر إلى مئة وفق الإصدار البرمجي weighted_heuristic_v1. وتشمل العوامل عمر الانتظار، ومدى إلحاح الإجراء المرتبط بالحالة، والنقص في عدد المحكّمين المطلوب، والدعوات أو التقارير المتأخرة، وعدد جولات التعديل.

تُحفظ أوزان هذه العوامل وحدود التطبيع في إعدادات المجلة، وتتحقق الخدمة من وجود وزن موجب واحد على الأقل. تعيد النتيجة الدرجة الكلية إلى جانب مساهمة كل عامل ووصف سببه، مثل وجود تقرير متأخر أو نقص محكّم. ثم ترتب المخطوطات تنازلياً بالدرجة، وتستخدم وقت الإرسال ومعرّف السجل لكسر التعادل. بهذه الصورة يستطيع رئيس التحرير أو مدير القسم فهم سبب تقدم مخطوطة في قائمة المتابعة، مع بقاء القرار حول الإجراء المناسب بشرياً.

8.10.2 المؤشرات ولوحات المعلومات

تجمع EditorialAnalyticsService البيانات التشغيلية في لوحة رئيس التحرير على مستوى المجلة. تعرض اللوحة أعداد المخطوطات والمنشورات والقرارات النهائية، ونسب القبول والرفض ضمن الحالات التي صدر فيها قرار نهائي، ووسيط زمن الوصول إلى القرار، ووسيط زمن التحكيم. كما تعرض توزيع الحالات، واتجاه الإرسال والنشر خلال الأشهر الاثني عشر الأخيرة، والتوزيع حسب الأقسام والموضوعات، وأعداد الدعوات والتقارير المتأخرة.

نُفذت تعريفات المؤشرات داخل الخدمة حتى لا تختلف طريقة الحساب بين الواجهات. فزمن القرار، مثلاً، يحسب من وقت إرسال المخطوطة إلى وقت القرار النهائي عند توفره، بينما يمتد زمن التحكيم في النسخة الحالية من إرسال الدعوة إلى تسليم التقرير، لأن سجل الدعوة لا يحتفظ بزمن استجابة مستقل يصلح للفصل بين مرحلتي الانتظار والتحكيم الفعلي. وتعرض اللوحة أيضاً أعلى عناصر قائمة الأولوية مع أسبابها، فتجمع بين الرؤية الإجمالية وإمكان الانتقال إلى المخطوطات التي تحتاج تدخلاً.

في الواجهة الأمامية تُعرض الملخصات في بطاقات، والتوزيعات والاتجاهات في رسوم وجداول مناسبة، مع حالات واضحة للتحميل وغياب البيانات والخطأ. ولا تُخزن هذه القيم على أنها قرارات ثابتة؛ بل تُشتق من السجلات الحالية عند طلب اللوحة، مما يبقيها متسقة مع تقدم المخطوطات وتغير الأعداد المنشورة.

> [صورة مقترحة: لوحة التحليلات التحريرية لرئيس التحرير، وتظهر المؤشرات الأساسية واتجاه الإرسال والنشر وتوزيع الحالات وقائمة المخطوطات الأعلى أولوية]

8.11 تنفيذ نظام كشف الانتحال

نُفذت منظومة الكشف عن التشابه المحتمل ضمن المشروع نفسه، وتكاملت مع النظام من خلال تطبيق integrity ومكتبة مستقلة لخط المعالجة. يفصل هذا التنظيم بين إدارة طلبات الفحص وحالاتها وصلاحياتها داخل Django، وبين معالجة الوثائق والفهرسة والاسترجاع والتحقق وبناء التقرير. وتُعامل النتيجة في جميع الطبقات على أنها دليل تشابه محتمل يحتاج إلى مراجعة بشرية، ولا يغير الفحص حالة المخطوطة ولا يصدر قرار رفض آلياً.

يمثل PlagiarismScreening عملية فحص لإصدار محدد، ويحفظ حالة الانتظار أو التنفيذ أو الاكتمال أو الفشل، وهوية طالب الفحص ومعرّف مهمة Celery، ونوع مجموعة المصادر وإصدارات خط المعالجة ونموذج التحقق، وملخص النتيجة والتقرير الكامل بصيغة JSON. يمنع قيد قاعدة البيانات وجود أكثر من فحص نشط للإصدار نفسه. ويُطلب الفحص تلقائياً بعد إنشاء الإصدار العربي الأول إذا كانت المنظومة مفعلة، كما يستطيع مدير القسم المسؤول إعادة طلبه يدوياً ضمن مرحلة الفحص الأولي. وقبل إنشاء الطلب تتحقق الخدمة من اللغة والحالة ووجود الملف وتهيئة المصادر والنموذج.

تُرسل العملية بعد تثبيت المعاملة إلى طابور Celery مخصص باسم plagiarism. يستهلكها عامل منفصل عن المهام العامة، ويثبت ملكية السجل قبل البدء لتجنب معالجة الطلب مرتين. ينزل العامل النسخة الكاملة من MinIO إلى ملف مؤقت، وينشئ كاشفاً من الإعدادات والموارد المحلية، ثم يحدث حالة الفحص ويحفظ النتيجة أو رسالة خطأ آمنة. وقد ضُبطت المهمة لتأكيد الرسالة بعد التنفيذ والسماح بإعادتها عند فقد العامل، مع بقاء الانتقال بين الحالات محكوماً بسجل قاعدة البيانات.

تبدأ المعالجة باستخراج النص من ملف PDF النصي، ثم تمريره إلى مكوّن معالجة مسبقة مشترك بين المخطوطة ومجموعة المصادر. يحتفظ المكوّن بالنص الأصلي المستخدم في عرض الأدلة، وبنسخة مطبّعة للاسترجاع، مع إزاحات البداية والنهاية التي تعيد كل مقطع إلى موضعه في الوثيقة. ويشمل التطبيع معالجة الفروق الكتابية العربية غير الدلالية وتقسيم النص إلى مقاطع تراعي الفقرات وأطوال النوافذ. أما ملفات PDF الممسوحة ضوئياً فلا تُجرى لها حالياً مرحلة تعرف ضوئي؛ فإذا تعذر استخراج نص كافٍ يفشل الفحص بصورة مضبوطة مع رسالة يمكن للمحرر فهمها.

تُجهز مجموعة المصادر خارج زمن فحص المخطوطة. يحفظ مستودع PostgreSQL الوثائق المصدرية ومقاطعها ونصوصها الأصلية والمطبّعة وإزاحاتها وبصمات المحتوى، مما يتيح منع التكرار وتتبع كل دليل إلى مصدره. يبني مسار الاسترجاع المعجمي فهرس BM25 محفوظاً على القرص، بينما تُحسب تمثيلات E5 الدلالية للمقاطع مسبقاً وتحفظ إما في pgvector مع فهرس متجهي أو في مخزن محلي من مصفوفات NumPy وفق إعداد التشغيل. وتبقى ملفات النماذج والفهارس والذاكرات الوسيطة موارد تشغيلية مستقلة عن شيفرة المستودع.

عند فحص مخطوطة، يسترجع كل مقطع مرشحين من المسار المعجمي ومسار E5 الدلالي. يدمج CandidateMerger القائمتين ويزيل الأزواج المكررة مع الاحتفاظ بإشارات المسارين، ثم تُرسل الأزواج المحدودة إلى نموذج تحقق محلي Siamese مبني على AraT5 ومحمل من نقطة التدريب الخاصة بالمشروع. يعمل النموذج على نصي مقطع المخطوطة والمصدر ويعيد احتمال التشابه الذي يستخدمه خط المعالجة لقبول الدليل أو استبعاده. يفصل هذا التسلسل بين مرحلة واسعة وسريعة للعثور على المصادر المحتملة ومرحلة أدق للتحقق من الأزواج المسترجعة.

بعد التحقق يتولى PlagiarismReportBuilder تنظيم الأدلة التي اجتازت العتبة في مخطط التقرير plagiarism_report_v1. يزيل الأدلة المتكررة، ويجمعها بحسب مقاطع المخطوطة ومصادرها، ويحسب تغطية المقاطع المشار إليها وأعلى احتمالاتها، وينشئ ملخصات على مستوى كل مصدر. ويحتوي كل دليل على النصوص اللازمة للمراجعة، وهوية المصدر، والإزاحات والدرجات ذات الصلة، بحيث يستطيع مدير القسم الانتقال من الملخص إلى المقاطع المقابلة بدلاً من الاعتماد على نسبة كلية غير مفسرة. وتظهر النتيجة بوصفها «تشابهاً محتملاً» مع علامة صريحة بضرورة المراجعة البشرية.

تعرض واجهة الفحص سجل العمليات للإصدار وحالة كل عملية وملخصها. ويقتصر الوصول إلى التقرير الكامل والأدلة التفصيلية على المستخدم التحريري المسؤول، بينما لا تُكشف تفاصيل داخلية حساسة أو مسارات ملفات أو رسائل استثناء خام. كما تُحفظ إصدارات المكونات ونوع قاعدة المصادر مع التقرير، الأمر الذي يتيح معرفة البيئة التي أنتجته عند المقارنة أو التقييم لاحقاً.

> [مخطط مقترح: التكامل التنفيذي بين Django وMinIO وRedis وCelery وعامل كشف الانتحال، ثم مراحل المعالجة من استخراج النص إلى الاسترجاع الهجين والتحقق وبناء التقرير]
>
> [صورة مقترحة: تقرير الفحص داخل مساحة مدير القسم، ويظهر ملخص التشابه المحتمل والمصادر والمقاطع المتقابلة وحالة ضرورة المراجعة البشرية]

8.12 تنفيذ نظام الإشعارات

نُفذت الإشعارات الحالية على شكل رسائل بريد إلكتروني غير متزامنة، ولم تُضف في النسخة الحالية علبة إشعارات داخلية مستقلة. تغطي الرسائل الأحداث التي تتطلب تنبيهاً عملياً، ومنها دعوة المحكّم، وإسناد المخطوطة أو إعادة إسنادها، وقرار الفحص المكتبي أو القرار التحريري، ورفع إصدار منقح للمحكّمين المرحّلين، والتذكير بموعد تقرير التحكيم، ونشر المقالة. تطلق خدمات المجال الرسالة بعد نجاح المعاملة باستخدام transaction.on_commit()، ثم تنفذها مهمة Celery. لذلك لا يؤدي فشل خادم البريد إلى التراجع عن قرار تحريري أو فقدان إصدار، ولا تُرسل رسالة لعملية أُلغيت معاملتها. وتعيد مهام البريد المحاولة عند أخطاء الاتصال المؤقتة بحد أقصى مضبوط وتأخير متزايد، بينما تستخدم مهمة التذكير قفلاً ووقت آخر إرسال للحد من الرسائل المكررة.

تُبنى الرسائل من سياق الحدث ومن بيانات المستخدم والمخطوطة، وتوجه القرارات الخاصة بالمؤلف إلى المؤلف المراسل في التنفيذ الحالي. ويستخدم إعداد التطوير مخرج البريد في الطرفية لتسهيل الفحص المحلي، بينما تُمرر بيانات خادم SMTP من متغيرات البيئة عند التشغيل في بيئة فعلية.

8.13 تنفيذ واجهات المستخدم

نُفذت الواجهة الأمامية باستخدام Next.js مع TypeScript، ونُظمت الصفحات وفق الدور وسياق العمل. يتعامل عميل Axios مع واجهات REST وإرفاق رمز المصادقة، وتستخدم TanStack Query لجلب بيانات الخادم وتخزينها المؤقت وإبطالها بعد العمليات التي تغير الحالة. أما النماذج فتستخدم React Hook Form مع مخططات Zod للتحقق المبكر من المدخلات، مع إعادة أخطاء الخادم إلى الحقول أو الرسائل العامة عند الحاجة.

تضم مساحة رئيس التحرير لوحة التحليلات وإدارة الأقسام وطلبات المحكّمين والأعداد وسجلات النشر. وتضم مساحة مدير القسم قوائم الفحص الأولي، وطلبات المحكّمين، ومراقبة سير المخطوطات، وأدوات الذكاء المساندة، ومساحة النشر ضمن نطاق القسم. أما محرر القسم فيعمل من مساحة المخطوطات المسندة إليه، حيث يراجع الإصدارات والتقارير ويستعرض توصيات المحكّمين ويرسل الدعوات ويتخذ القرار. ويملك المحكّم صفحة للدعوات والتكليفات وتقديم التقرير، بينما يستطيع المؤلف إنشاء الإرسال ومتابعة حالته وإصداراته وقراراته ورفع المراجعة المطلوبة.

رُبطت عناصر الواجهة بالصلاحيات وحالات المجال، فلا يظهر زر إجراء لمجرد أن الصفحة مفتوحة، بل يتأثر بدور المستخدم وعلاقته بالسجل وحالة المخطوطة. كما تعيد الواجهة التحقق من الخادم عند التنفيذ، لأن إخفاء العنصر في المتصفح لا يعد بديلاً عن التفويض الخلفي. وتعرض المكونات حالات التحميل والفراغ والفشل، وتعيد تحديث القوائم بعد الإسناد أو الدعوة أو القرار أو النشر حتى لا تبقى بيانات قديمة أمام المستخدم.

تعمل الواجهة العامة من دون اشتراط إنشاء حساب لعرض الصفحات المنشورة والبحث في المقالات وتصفح الأقسام والأعداد وتنزيل المقالات المنشورة وتصدير بياناتها الوصفية. أما المخطوطات غير المنشورة وملفات التحكيم ومساحات الإدارة فتظل خلف المصادقة والتحقق من الصلاحيات. وتعكس هذه الحدود الفرق بين بوابة الوصول العلمي العامة ومنطقة إدارة دورة التحرير.

الفصل التاسع

اختبارات النظام

يوضح هذا الفصل الاختبارات التي تم تطبيقها للتأكد من تلبية النظام للمتطلبات

9.1 مقدمة

يهدف اختبار النظام إلى التحقق من صحة الوظائف الأساسية، وسلامة التكامل بين مكوّناته، وفاعلية القيود الأمنية، وجودة المكوّنات الذكية، ومدى جاهزيته للعمل تحت أحمال متوقعة. وبالنظر إلى اتساع نطاق النظام وتعدد أدواره ومساراته التحريرية، اعتُمدت استراتيجية اختبار متعددة المستويات تجمع بين اختبارات الوحدات، واختبارات التكامل وواجهات البرمجة، واختبارات سير العمل، والتقييم التجريبي للمكوّن الذكي الخاص بكشف الانتحال. كما يميّز هذا الفصل بصورة صريحة بين النتائج التي نُفّذت وقِيست فعلياً، وحالات الاختبار الموجودة في الشفرة، والاختبارات المخططة التي لم يكتمل تنفيذها قبل إعداد التقرير. ويضمن هذا الفصل أن تكون الاستنتاجات مرتبطة بأدلة قابلة للتتبع، دون مساواة وجود اختبار في المستودع بنجاحه في آخر تشغيل.

9.2 استراتيجية الاختبار ونطاق التحقق

بُنيت استراتيجية التحقق حول المخاطر الأكثر تأثيراً في نظام إدارة مجلة علمية: سلامة انتقال المخطوطة بين الحالات، منع الوصول غير المصرح به، حماية سرية التحكيم، الحفاظ على تاريخ النسخ والمراجعات، صحة النشر والتصدير المعياري، واستقرار تنفيذ المهام غير المتزامنة. وتختلف طبيعة التحقق في المكوّنات الذكية؛ إذ لا يكفي التأكد من أن الواجهة البرمجية تعيد استجابة صحيحة، بل يجب تقييم جودة الاسترجاع والتصنيف والتجميع باستخدام مقاييس ملائمة للبيانات غير المتوازنة.

الجدول (9–1): تصنيف أدلة الاختبار المستخدمة

فئة الدليل	المقصود بها	طريقة عرضها في الفصل
نتيجة منفّذة ومقاسة	تشغيل أو تجربة محفوظة لها أرقام أو ناتج واضح	تُعرض بوصفها نتيجة فعلية مع نطاقها وقيودها

فئة الدليل	المقصود بها	طريقة عرضها في الفصل
تغطية آلية موجودة	حالة اختبار موجودة في الشفرة على النسخة المفحوصة	تُعرض بوصفها تغطية متاحة، لا بوصفها نجاحاً حديثاً
تحقق مخطط	سيناريو ومعايير قبول لم يكتمل تشغيلهما	يُعرض كخطة لاحقة ولا يدخل في نسب النجاح

يشمل النطاق الوظيفي الذي استهدفته الاختبارات إدارة الحسابات والأدوار، وإرسال المخطوطات وإصداراتها، والفرز الأولي، وتعيين المحررين والمحكّمين، والاستجابة للدعوات، وجولات المراجعة والقرارات، والنشر والتصدير، والاكتشاف عبر OAI-PMH، والإشعارات، والتخصيص، والميزات الذكية. أما اختبار الاختراق الشامل، واختبارات الحمل والإجهاد والتحمّل الطويل، والتحقق النهائي من جميع واجهات المستخدم في بيئة إنتاجية مماثلة، فلم يكتمل قبل إعداد هذا الفصل، ولذلك عوملت بوصفها حدوداً معلنة وخطة تحقق لاحقة.

لقطة التتبع: أُجري حصر مجموعة الاختبارات على النسخة ذات المعرّف 93ad1ff1e1cf7aa36f5fe6626666b60a2559488c. بلغ الحصر 322 حالة اختبار آلية في 21 ملفاً ضمن الخلفية. نجح أيضاً فحص التجميع النحوي لجميع ملفات Python في مجلد backend على هذه النسخة. ولا يعني ذلك أن مجموعة الحالات البالغ عددها 322 شُغّلت كاملة على النسخة نفسها.

9.3 اختبارات الوحدات والمكوّنات البرمجية

تستهدف اختبارات الوحدات أصغر القواعد القابلة للعزل، مثل التحقق من الانتقالات القانونية، وحساب أولوية المعالجة، وبناء البيانات الوصفية، وتقييم توصية المحكّمين، والتحقق من مخطط تقرير الانتحال. أما ملفات الاختبار التي تستخدم قاعدة بيانات Django وعميل واجهة البرمجة فتجمع غالباً بين اختبار الوحدة واختبار التكامل؛ لذلك يعرض الجدول الآتي توزيع الحالات بحسب النطاق البرمجي، وليس تصنيفاً صارماً لكل حالة منفردة.

الجدول (9-2): توزيع حالات الاختبار الآلية الموجودة في الخلفية

النطاق البرمجي	عدد الحالات	أمثلة على السلوك المغطى
الحسابات وطلبات الانضمام للتحكيم	33	التفعيل، الملف الشخصي، ORCID، طلبات المحكّمين والموافقة عليها
الخدمات الأساسية والتوصية	6	التخزين، الروابط المؤقتة، درجات التوصية وترتيبها
الاكتشاف العلمي OAI-PMH	17	الأفعال القياسية، المرشحات، الأخطاء، والبيانات العربية
كشف الانتحال والتكامل معه	32	النماذج، الخدمة، المهام، الواجهة، والربط التلقائي بالإرسال
المجلة والتحليلات	31	الأقسام، الإعدادات، الموضوعات، لوحة رئيس التحرير والخصوصية
الإشعارات	8	طمس الهوية، القرارات، التذكير بالمواعيد ومنع التكرار
النشر والتخصيص العام	83	المسودات، النشر، التصدير، التنزيل، الأذونات وصفحات المحتوى
التحكيم والتوصية بالمحكّمين	44	الدعوات، المهل، التزامن، الاستبدال، السرية وجولات المراجعة
الإرسال والإصدارات	33	الإنشاء، الملفات، لوحة المؤلف، الخصوصية ورفع النسخ المعدلة
سير العمل والأولوية	34	التعيين، الطوابير، الانتقالات، الترتيب ومسار متكامل واحد
عقد عناوين API	1	الحفاظ على مسارات الواجهات العامة
المجموع	322	حالات موجودة في الشفرة على النسخة المفحوصة

تُظهر هذه التغطية اهتماماً واضحاً بقواعد المجال، ولا سيما الحالات السلبية التي تمنع إجراءً غير صالح أو وصولاً غير مخوّل. ومن أمثلتها رفض إنشاء مسودة نشر من مخطوطة غير مقبولة، ومنع المؤلف من الوصول إلى مخطوطة مؤلف آخر، ورفض محكّم من خارج القسم، ومنع تعديل حالة المقال مباشرة، والتحقق من عدم تجاوز عدد المحكّمين المقبولين عند الاستجابات المتزامنة. وتعد هذه الحالات أكثر قيمة من الاقتصار على المسار الناجح، لأنها تختبر القيود التي تحافظ على اتساق النظام وسرية العملية التحريرية.

9.3.1 نتائج التشغيل الجزئي الموثقة أثناء التطوير

توجد نتائج تشغيل جزئية موثقة من مراحل سابقة من التطوير لثلاث حزم، كما يبين الجدول التالي. كانت هذه النتائج ناجحة في وقت تشغيلها، إلا أن عدد الحالات في النسخة الحالية ازداد لاحقاً؛ لذلك لا تُستخدم لإثبات نجاح جميع الحالات الحالية، وإنما لإثبات أن أجزاء رئيسية شُغّلت فعلياً خلال دورة التطوير.

الجدول (9-3): نتائج تشغيل جزئية موثقة مقارنة بحجم الحزم الحالية

الحزمة	الحالات المنفذة آنذاك	النتيجة الموثقة	الحالات الموجودة حالياً
الاكتشاف OAI-PMH	16	نجحت	17
النشر	59	نجحت	83
المجلة والأقسام	21	نجحت	26
المجموع الجزئي	96	نجحت في لقطات سابقة	126

بلغ عدد الحالات الناجحة في هذه التشغيلات الجزئية 96 حالة. ومع ذلك، فإن الحالات الثلاثين المضافة لاحقاً إلى الحزم نفسها، فضلاً عن بقية الحزم، تحتاج إلى تشغيل موحد على النسخة النهائية قبل الادعاء بنجاح المجموعة الكاملة. ويمثل غياب هذا التشغيل الشامل أحد أهم قيود التحقق في النسخة المقدمة.

9.4 اختبارات التكامل وسير العمل

تتحقق اختبارات التكامل من تعاون النماذج والخدمات وواجهات البرمجة والأذونات والمهام الخلفية، مع التركيز على الحدود التي تنتقل عندها البيانات بين المكوّنات. وقد استُخدم عميل واجهات Django في عدد كبير من الحالات لإرسال طلبات مصادق عليها والتحقق من رمز الاستجابة ومحتوى الاستجابة والتغييرات المخزنة في قاعدة البيانات. وتغطي الاختبارات كذلك السلوك التعويضي، مثل عدم إفشال إرسال المخطوطة عند تعذر جدولة فحص الانتحال، وتسجيل رسالة آمنة عند فشل خط المعالجة.

الجدول (9-4): سيناريوهات تكامل ممثلة في مجموعة الاختبار

السيناريو	المكوّنات المتكاملة	معيار التحقق
إرسال مخطوطة عربية	الإرسال، التخزين، نسخة المخطوطة، خدمة الانتحال والمعاملة	إنشاء الإرسال ثم تسجيل الفحص بعد نجاح المعاملة دون تعطيل الإرسال
الفرز وتعيين محرر	قائمة المدير، التقييم الأولي، الأذونات وخدمة التعيين	منع التعيين قبل اكتمال الفرز وحصر المحررين في القسم الصحيح
دعوات المحكّمين	البحث، التوصية، الدعوات، المهل والتزامن	قبول العدد المطلوب وإنهاء الدعوات الزائدة دون تجاوز السعة
جولة تعديل ثم قبول	المراجعات، قرار المحرر، نسخة معدلة، جولة جديدة والنشر	الحفاظ على التاريخ والوصول إلى مسودة نشر من النسخة المقبولة الأخيرة
نشر مقال واكتشافه	النشر، العدد، البيانات الوصفية، التصدير وOAI-PMH	إظهار المنشور فقط وتوليد BibTeX وRIS وDublin Core وOAI-DC صحيح
تنفيذ فحص الانتحال	Celery، الخدمة، خط المعالجة، التقرير والواجهة	حفظ تقرير مطابق للمخطط أو تسجيل فشل آمن قابل للمراجعة

تتضمن المجموعة الحالية اختباراً آلياً لمسار تحريري متكامل يبدأ من تقديم مراجعة، ثم طلب تعديل، ورفع نسخة جديدة، واتخاذ قرار بالقبول، وانتهاء بإنشاء مسودة نشر. يمثل هذا الاختبار دليلاً جيداً على قابلية صياغة المسار المتكامل آلياً، إلا أن عدم توفر ناتج تشغيل نهائي موحد يمنع اعتباره دليلاً حديثاً على نجاح المسار في النسخة الحالية. وينطبق الأمر نفسه على التكامل الفعلي مع PostgreSQL وRedis وMinIO وعمال Celery في بيئة تشغيل كاملة؛ فاختبارات عديدة تستخدم العزل والمحاكاة، وهي مفيدة لتحديد الخطأ، لكنها لا تستبدل اختبار قبول في بيئة خدمات متكاملة.

9.5 التحقق من الجوانب الأمنية

تُعد السرية والتحكم بالوصول من المتطلبات الجوهرية في نظام يدير مخطوطات غير منشورة وتحكيماً مزدوج التعمية. وقد غطت الحالات الآلية الحالية عدداً من الضوابط الأمنية على المستوى الوظيفي. ولا تُعد هذه الحالات اختبار اختراق، لأنها لا تبحث

منهجياً عن ثغرات غير معروفة ولا تستخدم فاحصاً ديناميكياً مستقلاً؛ لكنها تتحقق من أن قواعد الأذونات والخصوصية المقصودة مطبقة في نقاط حساسة.

الجدول (9–5): التغطية الأمنية الوظيفية وحدودها

المحور الأمني	أمثلة على التحقق الموجود	حد التغطية
المصادقة ودورة الحساب	رفض دخول المستخدم غير النشط وإبطال رمز الوصول بعد تعطيل الحساب	لا يشمل اختبار سرقة الرموز أو معدل محاولات الدخول
التحكم القائم على الأدوار	منع الأدوار غير المخولة وحصر مدير القسم في قسمه والمحرر في المخطوطات المسندة إليه	تغطية وظيفية وليست مراجعة شاملة لكل مسار
سرية التحكيم	توفير الملف المعمّى للمحكّم المقبول ومنع كشف الملف الكامل أو هوية المحكّم	لم يُنفّذ تدقيق يدوي شامل لتدفق البيانات في الواجهة
تقليل كشف البيانات	إخفاء مسارات الملفات والحقول الداخلية والمسودات والمقالات المسحوبة وإعدادات الترتيب الداخلية	لا توجد أداة منع تسرب بيانات مستقلة
التحقق من المدخلات والأخطاء	رفض ملفات غير صالحة وإرجاع رسائل نظيفة عند فشل التخزين أو خط الانتحال	لم يُنفّذ فحص DAST أو اختبار حقن منهجي

بناءً على ذلك، يمكن القول إن النظام يمتلك تغطية آلية جيدة للمخاطر الأمنية المرتبطة بمنطق الصلاحيات وخصوصية البيانات، مع بقاء الحاجة إلى مراجعة إعدادات النشر الفعلية، وسياسات CORS وCSRF، وصلاحيات حاويات MinIO، وإدارة الأسرار، ومحددات معدل الطلبات، إضافة إلى فحص SAST وDAST واختبار اختراق محدود قبل النشر العام.

9.6 تقييم المكوّنات الذكية

9.6.1 مقاييس التقييم

يختلف تقييم المكوّنات الذكية عن اختبار الوظائف التقليدية. ففي بيانات كشف الانتحال تكون الأزواج الإيجابية أقل بكثير من السلبية، ولذلك قد تعطي الدقة العامة Accuracy قيمة مرتفعة حتى لنموذج يميل إلى التنبؤ السلبي. لهذا استُخدمت

الدقة الإيجابية Precision والاستدعاء Recall ودرجتا F1 وF2 ومتوسط الدقة AP، مع فصل تقييم الاسترجاع عن تقييم التحقق وتجميع المقاطع.

$$\text{Precision} = TP / (TP + FP)$$

$$\text{Recall} = TP / (TP + FN)$$

$$F1 = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

تقيس Precision نسبة التنبيهات الصحيحة بين التنبيهات الصادرة، بينما يقيس Recall نسبة حالات الانتحال المرجعية التي استُعيدت. وتوازن F1 بينهما، في حين تمنح F2 وزناً أكبر للاستدعاء، وهو أمر مهم في الفحص الأولي حيث تكون خسارة المصدر الحقيقي صعبة التعويض لاحقاً. ومع ذلك، لا تبرر هذه المقاييس اتخاذ قرار تأديبي تلقائي؛ إذ يبقى التقرير أداة دعم قرار تتطلب مراجعة بشرية.

9.6.2 تقييم الاسترجاع الهجين

قُيّم الاسترجاع الهجين القائم على BM25 والتمثيل الدلالي متعدد اللغات على 438 وثيقة مشتبه بها تضم 1,725 إشارة مرجعية. شمل التقييم 10,248 مقطعاً مستعلماً و188,787 زوجاً مرشحاً. واعتمد المقياس الصارم لاستدعاء مقطع المصدر عند الرتبة K على تحقق شرطين معاً: الوصول إلى وثيقة المصدر الصحيحة، والتداخل مع المقطع المرجعي داخلها.

الجدول (9-6): نتائج الاسترجاع الهجين حسب عمق الاسترجاع

K	استدعاء وثيقة المصدر	استدعاء مقطع المصدر	عدد إصابات المقاطع
1	0.9507	0.9455	1,631
3	0.9780	0.9716	1,676
5	0.9832	0.9768	1,685
10	0.9896	0.9826	1,695

بلغ استدعاء مقطع المصدر 98.26% عند K=10، ما يدل على أن مرحلة الاسترجاع تضع المصدر الصحيح ضمن المرشحين في معظم الحالات. في المقابل، يزداد عدد الأزواج التي يجب أن يعالجها نموذج التحقق مع رفع K، وهو ما يرفع الكلفة واحتمال الإيجابيات الكاذبة. لذلك لا تكفي نتيجة الاسترجاع وحدها للحكم على جودة النظام النهائي، بل يجب تفسيرها مع نتائج نموذج التحقق وسياسة اختيار المرشحين.

9.6.3 الاختبار الرسمي لنموذج التحقق

ثُبّتت نقطة الحفظ من الحقبة السابعة والعتبة 0.349609375 قبل تشغيل الاختبار الرسمي على 601 وثيقة و254,915 زوجاً مرشحاً. عولجت الأزواج على دفعات حجمها 8 باستخدام BF16، وسُجلت المخرجات تدريجياً مع إمكانية استئناف التنفيذ. يلخص الجدول التالي المقاييس الرسمية على مستوى أزواج المقاطع.

الجدول (9-7): نتائج الاختبار الرسمي لنموذج التحقق على مستوى الأزواج

المقياس	القيمة
متوسط الدقة AP	0.424641
Precision	0.297006
Recall	0.557048
F1	0.387438
F2	0.474039
TP / FP / TN / FN	5,268 / 12,469 / 232,989 / 4,189

أظهر الاختبار الرسمي استدعاءً متوسطاً مع Precision منخفضة نسبياً؛ إذ قُبل 17,737 زوجاً، وكان نحو 29.7% منها صحيحاً. تعكس هذه النتيجة صعوبة المهمة وشدة عدم توازن البيانات، كما تفسر ضرورة وجود مرحلة تجميع وسياسة تشغيل تقللان الضجيج قبل إنشاء التقرير التحريري. وينبغي عدم مقارنة F1 في هذا الاختبار مباشرة بنتائج دراسات تستخدم تصنيف أزواج وثائق كاملة أو تقسيمات مختلفة، لأن وحدة القياس والبروتوكول ليسا متطابقين.

9.6.4 تقييم التجميع وسياسة التشغيل

قُيّمت 16 قيمة لفجوات تجميع المقاطع على جانبي المخطوطة والمصدر. ثم أُجري مسح مشترك لقيم K وعتبة نموذج التحقق مع تثبيت منطق التجميع. اختيرت على بيانات التحقق سياسة Top-1 مع عتبة 0.5 وفجوتي تجميع 300/300، فحققت تحسناً كبيراً في الدقة النهائية وخفضت حجم العمل المطلوب من نموذج التحقق.

الجدول (9-8): أثر سياسة اختيار المرشحين في النتائج النهائية على بيانات التحقق

الإعداد	Precision النهائية	Recall النهائي	F1 النهائي	الأزواج المحتفظ بها
جميع المرشحين، عتبة 0.349609375	0.176955	0.829582	0.291690	33,165
Top-5، العتبة نفسها	0.321519	0.816720	0.461399	8,950
Top-1، عتبة 0.5	0.858182	0.758842	0.805461	1,790

خفضت السياسة المختارة عدد الأزواج المحتفظ بها بنسبة 94.60% مقارنة بالاتحاد الكامل، وحققت F1 مقدارها 0.805461 على بيانات التحقق. إلا أن هذه النتيجة اختيرت على مجموعة التحقق نفسها، ولم يُثبت تشغيلها لاحقاً على مجموعة اختبار جديدة غير مستخدمة في الاختيار؛ ولذلك تُعرض بوصفها نتيجة اختيار إعداد تشغيل، لا نتيجة اختبار نهائي مستقلة. أما النتيجة الرسمية غير المتحيزة المتاحة فتبقى نتيجة نموذج التحقق على مستوى الأزواج في الجدول (9-7).

9.6.5 بقية الميزات الذكية

توجد اختبارات وظيفية لمنطق توصية المحكّمين تراجع إظهار دليل التشابه الدلالي وتداخل الكلمات المفتاحية وترتيب النتائج ثم الحمل الحالي، كما يوجد اختبار يثبت أن المحكّم المعتمد يظهر في توصيات القسم المناسب فقط. وتغطي اختبارات نمذجة الموضوعات تجميع الموضوعات ضمن القسم وصلاحيات الاطلاع وعدم كشف بيانات خاصة. أما ترتيب الأولويات فتغطيه حالات لحساب عمر الانتظار وتفسير المواعيد المتأخرة وترتيب المخطوطات حسب الدرجة.

الجدول (9-9): حالة تقييم الميزات الذكية الأخرى

الميزة	التحقق المتاح	التقييم غير المكتمل
توصية المحكّمين	اختبارات درجات التشابه والكلمات المفتاحية والحمل والأذونات	لا توجد مجموعة موسومة لقياس Precision@K أو NDCG
نمذجة الموضوعات	اختبارات التجميع الإحصائي والعرض والصلاحيات	لا يوجد تقييم موضوعي للاتساق Coherence أو تقييم بشري للعناقيد

الميزة	التحقق المتاح	التقييم غير المكتمل
ترتيب الأولويات	اختبارات حساب الدرجة والتفسير والترتيب	لا يوجد تقييم ميداني لمدى توافق الترتيب مع قرارات هيئة التحرير

تثبت الاختبارات الحالية صحة الربط الوظيفي والمنطق الحسابي لهذه الميزات، لكنها لا تكفي لادعاء جودة تنبؤية أو توصوية مرتفعة. يتطلب ذلك بيانات مرجعية يحدد فيها خبراء المجال المحكّمين الأنسب والموضوعات الصحيحة وترتيب الأولويات المقبول، ثم حساب مقاييس مناسبة ومقارنة خطوط أساس متعددة.

9.7 اختبارات الأداء والاعتمادية

9.7.1 مشاهدات الأداء المتاحة

لم يُنفّذ اختبار حمل شامل على النظام المتكامل، إلا أن تجارب مكوّن كشف الانتحال وفرت بعض القياسات الجزئية التي تساعد في فهم مواضع الكلفة. وهذه القياسات تخص تشغيلات ومكونات محددة، ولا تمثل زمن استجابة واجهة النظام تحت مستخدمين متزامنين.

الجدول (9–10): قياسات أداء جزئية متاحة لمكوّن كشف الانتحال

المشاهدة	القيمة والسياق	الدلالة
تضمين مقاطع مخطوطة	40 مقطعاً في 10 دفعات على CPU خلال نحو 22 ثانية	التضمين الدلالي جزء ملحوظ من زمن الفحص
بحث HNSW البعيد	بين 0.3 و0.5 ثانية للمقطع بعد إنشاء الفهرس في أمثلة التشغيل	تراكم الاستعلامات يزيد الزمن ويتأثر بالشبكة
الاختبار الرسمي	معالجة 254,915 زوجاً بدفعات حجمها 8 وBF16	يثبت إمكان إكمال استدلال كبير، لكن الزمن الكلي غير محفوظ
سياسة Top-1	خفض الأزواج 94.60% على بيانات التحقق	تحسين محتمل للكلفة مع تضحية محسوبة بالاستدعاء

تؤكد هذه المشاهدات أن عدد المرشحين وسياسة التخزين المؤقت وعدد عمال التحليل عوامل حاسمة. كما أن تحميل النماذج والفهارس عند كل طلب غير مقبول؛ ولذلك يعتمد التصميم على نسخة مخبأة من خط المعالجة داخل العامل طويل العمر، مع فصل طابور الانتحال عن الطابور العام. ومع ذلك، بقيت ذروة الذاكرة، وزمن البداية الباردة، وزمن فحص مخطوطة كاملة، وسعة الطابور، والإنتاجية تحت التزامن غير مقاسة قياساً نهائياً.

9.7.2 خطة اختبارات الحمل والإجهاد والتحمّل

لإكمال التحقق، وُضعت خطة أولية قابلة للتنفيذ بأداة توليد حمل مثل k6. تفصل الخطة بين واجهات REST التفاعلية وبين فحص الانتحال غير المتزامن؛ إذ يُقاس زمن قبول مهمة الفحص بسرعة، بينما تُقاس مدة المعالجة وسعة الطابور كمقاييس خلفية مستقلة. القيم الآتية أهداف قبول مقترحة للتجربة الأولى، وليست نتائج تم قياسها.

الجدول (9-11): خطة اختبارات الأداء التي لم يكتمل تنفيذها

نوع الاختبار	ملف الحمل المقترح	المقاييس	معيار القبول الأولي
الحمل Load	20 مستخدماً افتراضياً لمدة 15 دقيقة على القراءة والبحث والطوابير	p95، معدل الأخطاء، الإنتاجية	p95 أقل من ثانيتين للقراءة ومعدل أخطاء أقل من 1%
الإجهاد Stress	زيادة تدريجية من 20 إلى 100 مستخدم خلال 20 دقيقة	نقطة الانكسار، أخطاء 5xx، اتساق البيانات	تدهور تدريجي دون فقد بيانات أو تجاوز غير صحيح للحالات
الاندفاع Spike	انتقال من 10 إلى 100 مستخدم لدقيقتين ثم العودة	الاستعادة، زمن الطابور، الاتصالات	عودة المؤشرات إلى خط الأساس وعدم ازدواج العمليات
التحمّل Soak	10 مستخدمين لمدة ساعتين مع مهام دورية	الذاكرة، الاتصالات، المهام المتراكمة	غياب نمو مستمر للذاكرة وعدم تراكم مهام غير منتهية

ينبغي تشغيل هذه الاختبارات على بيانات اختبار لا تحتوي مخطوطات حقيقية، مع تثبيت مواصفات العتاد وإصدارات الخدمات وحجم قاعدة البيانات. كما يجب مراقبة PostgreSQL وRedis وCelery وMinIO كل على حدة، لأن ارتفاع

زمن الاستجابة قد ينتج عن قاعدة البيانات أو التخزين أو الطوابير لا عن طبقة API نفسها. ويُستحسن تكرار كل تجربة ثلاث مرات والإبلاغ عن الوسيط والنطاق، مع حفظ سكربتات الحمل وملفات النتائج لتسهيل إعادة التجربة.