نموذج لإنشاء مستودع وطني للبيانات البحثية

د. محمد مسعد محمد الزلباني
مكتب مكتبة الكونجرس
mh_zalabany@hotmail.com

المستخلص

يناقش البحث القوانين والاتفاقيات المنظمة لإنشاء مستودع للبيانات البحثية ، سواء على المستوى الدولي أو المحلي، بإضافة إلى المعايير الأخلاقية والسياسات الجامعية التي ستنظم عمليات إدارة البيانات البحثية ومشاركتها. كما تناول بالتفصيل سياسات إدارة البيانات، بدءًا من جمع البيانات وتنظيمها وتخزينها، وصولًا إلى إتاحتها، ومراقبة الوصول إليها، وتأمينها، والنسخ الاحتياطي، وغيرها من الإجراءات اللازمة لضمان الحماية والحوكمة الفعالة للبيانات.

كما تطرق البحث أيضًا إلى البنية التحتية والتكنولوجية المطلوبة لإنشاء المستودع الوطني للبيانات البحثية، مسترشدًا بالبنية التحتية الخاصة بجامعة هارفارد في تطبيق منظومة داتافيرس (Dataverse).كما استعرض الأنظمة والتجهيزات البرمجية الخاصة بإدارة البيانات البحثية، بما في ذلك أنظمة المستودعات والبرمجيات المساندة، واحتياجاتها التشغيلية لضمان استمرارية عملها بكفاءة. إضافة إلى ذلك، ناقش البحث سياسات ضبط المستودع وتهيئته، ومعايير البيانات الوصفية المستخدمة، ثم تناول تقييم أداء المستودع من خلال وضع مؤشرات لقياس فعاليته وكفاءته.

وفي الختام، قدم البحث نموذجًا تطبيقيًا لإدارة المستودع الوطني للبيانات البحثية، يشمل التخطيط المالي، والتشغيلي، والإداري لضمان كفاءة عمل المستودع واستدامته، بما يحقق الأهداف المرجوة في دعم البحث العلمي وتعزيز مشاركة البيانات البحثية على المستويين المحلي والدولي.

الكلمات المفتاحية : مستودعات البيانات البحثية – حركة العلم المفتوح – خطة وطنية للبيانات البحثية.

تاريخ استلام البحث: 2025/9/7 تاريخ إجازة البحث: 2025/9/27

0 تمهيد

تعتبر المستودعات الوطنية للبيانات البحثية من العناصر الأساسية لتعزيز البحث العلمي ودعم الابتكار في مختلف المجالات الأكاديمية. حيث تلعب هذه المستودعات دورًا حيويًا في توفير بنية تحتية متكاملة لحفظ ونشر البيانات الناتجة عن الأبحاث العلمية على المستوى الوطني، مما يسهل إعادة استخدامها ويوسع نطاق تأثيرها. كما أن إنشاء نموذج مستودع وطني للبيانات البحثية في مصر يمثل خطوة استراتيجية لدعم جهود البحث العلمي على المستوى الوطني وتعزيز التعاون بين الباحثين من مختلف المؤسسات الأكاديمية والبحثية. ويعتبر هذا المستودع أداة أساسية لتنظيم وتخزين البيانات وفق معايير دقيقة، مما يسهل استخدام البيانات البحثية ويدعم حركة البيانات المفتوحة، بالإضافة إلى توفير وسيلة موثوقة لحفظ البيانات البحثية بشكل دائم، مما يعزز من شفافية البحث العلمي (Tenopir et al., 2011).

1 نموذج مقترح لإنشاء مستودع وطني للبيانات البحثية

تعتبر فكرة إنشاء مستودع وطني للبيانات البحثية خطوة استراتيجية مهمة لتعزيز البيئة العلمية في مصر وتحفيز البحث والابتكار. سيوفر هذا المستودع منصة مركزية لحفظ وإدارة البيانات البحثية بشكل آمن وفعال، مما يسهل على الباحثين الوصول إلى البيانات ومشاركتها. كما يعزز التعاون بين المؤسسات البحثية على المستويين الوطني والدولي من خلال إتاحة البيانات البحثية بشكل مفتوح ومنظم. بالإضافة إلى ذلك، سيساهم المستودع في تسريع وتيرة الاكتشافات العلمية وتطوير الحلول للتحديات الاجتماعية والاقتصادية. علاوة على ذلك، سيعزز المستودع الشفافية والنزاهة في البحث العلمي من خلال ضمان إمكانية إعادة استخدام البيانات والتحقق منها (Conzett et al., 2022).

سيساهم المستودع الوطني للبيانات البحثية في تعزيز مستوى البحث العلمي في مصر، وزيادة تواجد الباحثين المصريين في الساحة العلمية العالمية من خلال تحسين إمكانية الوصول إلى البيانات البحثية، مما يعزز سمعة الباحثين (Chigwada & Kasiroori, 2015). بعد استعراض بعض التجارب الوطنية في إنشاء مستودعات للبيانات البحثية على المستوى الوطني، أصبح من الضروري وضع استراتيجية شاملة كنموذج تطبيقي يمكن الاستفادة منه في حال الرغبة في إنشاء مستودع للبيانات البحثية، على غرار المشاريع الوطنية التي تم تناولها في بداية هذا الفصل. وقد توصلنا إلى أنه من أجل وضع هذه الاستراتيجية، يجب الالتزام بالمعايير الدولية.

وفيما يلي استراتيجية دولية ومجموعة من المتطلبات اللازمة لإنشاء مثل هذا المستودع، وعناصرها كالتالي:

1/1 - التخطيط الاستراتيجي والحوكمة

الاستراتيجية الوطنية:

تُعتبر صياغة استراتيجية وطنية للبيانات البحثية خطوة أساسية نحو تعزيز البيئة البحثية والعلمية في أي دولة. تهدف هذه الاستراتيجية إلى تحسين عمليات جمع وإدارة وتبادل البيانات البحثية بشكل مستدام وفعّال، مما يعزز من دور البحث العلمي في دعم الاقتصاد والمجتمع. وفيما يلي إطار عام يمكن الاعتماد عليه لوضع استراتيجية وطنية للبيانات البحثية، بما يتماشى مع الأطر الدولية مثل السحابة الأوروبية للعلوم المفتوحة (EOSC) (Guidance National Data Strategy, 2019).

أهداف الاستراتيجية:

- تعزيز الوصول المفتوح للبيانات البحثية وزيادة مشاركة البيانات.
- ضمان جودة وموثوقية البيانات البحثية.
- تسهيل التعاون بين الباحثين على المستويات المحلية والدولية.
- تعزيز الابتكار من خلال الاستفادة الفعالة من البيانات البحثية.
- دعم الشفافية في البحث العلمي.

النتائج المتوقعة

- تحسين جودة الأبحاث العلمية في الدولة وزيادة إنتاجيتها.
- تعزيز التعاون بين الباحثين المحليين والدوليين.
- توفير بيانات بحثية ذات قيمة لدعم الابتكار والتنمية الاقتصادية.
- رفع مستوى الشفافية والمساءلة في عمليات البحث العلمي.

إن تطبيق هذه الاستراتيجية من شأنه أن يضع الدولة في مكانة رائدة في مجال البحث العلمي ويعزز من دور البيانات في دعم التنمية المستدامة والمساهمة في تقدم المعرفة على المستوى الوطني والدولي(*Policy paper National Data Strategy Mission 1 Policy Framework: Unlocking the value of data across the economy*, 2021).

الحوكمة :

تشير الحوكمة إلى مجموعة من القواعد والإجراءات والهياكل التي تنظم وتوجه عملية اتخاذ القرارات داخل المؤسسات أو الهيئات. وتهدف الحوكمة إلى تحقيق الشفافية والمساءلة والكفاءة في إدارة الموارد واتخاذ القرارات. تُطبق الحوكمة في مجالات متعددة، مثل الحوكمة المؤسسية التي تركز على تنظيم العلاقة بين إدارة المؤسسة وأصحاب المصلحة، مثل المساهمين والموظفين والعملاء، لضمان تحقيق أهداف المؤسسة بطرق قانونية وأخلاقية.

لذا، من الضروري عند إنشاء مشروع ناجح لإدارة البيانات البحثية على المستوى الوطني أن نضع هيكل حوكمة واضح يتضمن أدوارًا ومسؤوليات محددة. يجب أن يشمل هذا الهيكل جميع أصحاب المصلحة من الحكومة والأوساط الأكاديمية والصناعة، بما في ذلك ممثلين عن المؤسسات البحثية ووكالات التمويل ومستخدمي البيانات، وهو ما سيتم تناوله بالتفصيل لاحقًا في هذا الفصل.

1/2- السياسات والامتثال

لإدارة مستودعات البيانات البحثية في مصر بطريقة قانونية وآمنة، يتعين الالتزام بمجموعة من القوانين واللوائح الدولية والمحلية التي تنظم عمليات جمع البيانات وتخزينها ومعالجتها ومشاركتها. فيما يلي أبرز القوانين والاتفاقيات الدولية ذات الصلة، بالإضافة إلى القوانين المحلية التي يجب أخذها بعين الاعتبار:

أولًا: القو انين والاتفاقيات الدولية

سياسة حماية البيانات والخصوصية : يجب وضع سياسة شاملة لحماية البيانات، مستندة إلى قوانين حماية الملكية الفكرية وحماية البيانات الشخصية، سواء كانت محلية إن وجدت أو عالمية تُطبق في العديد من الدول. تحدد هذه السياسة المبادئ الأساسية التي ينبغي اتباعها لضمان الحماية والخصوصية اللازمة للبيانات، سواء كانت شخصية أو غيرها من أنواع البيانات الموجودة في مستودعات البيانات البحثية.

في هذا السياق، أصدر القانون المصري قانون رقم 151 لسنة 2020 لحماية البيانات الشخصية، والذي تم تحديثه في عام 2023. كما يوجد أيضًا قانون حماية الملكية الفكرية (قانون حماية حقوق الملكية الفكرية رقم 82 لسنة 2002) الذي يختص بحماية المعلومات وضمان نسبتها لأصحابها("قانون حماية حقوق الملكية الفكرية رقم 82 لسنة 2002", ٢٠٢٠).

- **اللائحة العامة لحماية البيانات (GDPR)** تُعتبر اللائحة العامة لحماية البيانات (GDPR) التي أقرها الاتحاد الأوروبي من بين أكثر الأنظمة صرامة في مجال حماية البيانات الشخصية ولذا يوصي الباحث باتباعها في النموذج المصري المقترح. حيث تفرض هذه اللائحة قيودًا صارمة على جمع ومعالجة البيانات الشخصية للأفراد داخل الاتحاد الأوروبي، لذا، إذا كان المستودع سيحتوي على بيانات شخصية تخص مواطنين أو مقيمين في الاتحاد الأوروبي، فإنه يتعين الالتزام بهذه اللوائح (Mesarčík & Hamulak, 2024).

- متطلبات تنفيذ GDPR:
  - الحصول على موافقة صريحة من الأفراد قبل جمع بياناتهم.
  - حماية البيانات الشخصية من الاختراق أو التسرب.
  - السماح للأفراد بالوصول إلى بياناتهم والمطالبة بحذفها عند الطلب.

- **اتفاقية برن لحماية المصنفات الأدبية والفنية Berne Convention:** تُعتبر هذه الاتفاقية أول اتفاقية عالمية تهتم بحماية حقوق الملكية الفكرية، حيث يُنظر إلى "اتفاقية برن" على أنها الأساس القانوني لتنظيم حقوق المؤلف والحقوق المجاورة على الصعيد الدولي. ويُعزى ذلك إلى كونها واحدة من أوائل الاتفاقيات التي تم التوصل إليها لمعالجة قضايا حقوق المؤلف (بسام، 2005).

تهدف اتفاقية برن إلى حماية حقوق الملكية الفكرية، بما في ذلك البيانات والمواد البحثية التي قد تُعتبر ملكية فكرية. في حال احتوى المستودع على بيانات أو مواد بحثية تعود لمؤلفين أو باحثين دوليين، فإنه يتوجب حماية حقوقهم وفقًا لهذه الاتفاقية (Summary of the Berne Convention for the Protection of Literary and Artistic Works (1886).

- **متطلبات تطبيقها:**
  - الالتزام بحقوق النشر عند استخدام أو مشاركة البيانات.
  - حماية حقوق الملكية الفكرية للمساهمين في المستودع.

- **معايير FAIR:** تهدف معايير FAIR (قابلة للاكتشاف، قابلة للوصول، قابلة للتشغيل البيني، قابلة لإعادة الاستخدام) إلى تحسين جودة البيانات البحثية وسهولة الوصول

إليها. من المهم ملاحظة أن هذه المعايير ليست قوانين ملزمة، لكنها تحظى بتأييد واسع في الأوساط العلمية العالمية.

- **متطلبات تنفيذها:**
  - جعل البيانات قابلة للبحث والوصول إليها بشكل آمن.
  - ضمان قابلية تشغيل البيانات مع أنظمة أخرى.
  - الحفاظ على حقوق الوصول وإعادة الاستخدام بشكل منظم.

اتفاقيات منظمة التجارة العالمية (WTO) المتعلقة بتدفق البيانات عبر الحدود: تضمن هذه الاتفاقيات حرية تدفق البيانات عبر الحدود وفقًا لمعايير حماية محددة، وتعتبر ضرورية لنقل البيانات البحثية بين مصر ودول أخرى (Mishra). ويتم تطبيقها عند نقل البيانات البحثية خارج الحدود المصرية إلى دول أخرى.

- متطلبات تنفيذها:
  - ضمان حماية البيانات عند نقلها دوليًا.
  - التحقق من الامتثال للوائح حماية البيانات في الدول المستقبلة.

ثانيًا: القوانين المحلية في مصر

قانون حماية البيانات الشخصية المصري رقم ١٥١ لسنة ٢٠٢٠: يُعتبر هذا القانون الجديد أساساً لحماية البيانات الشخصية في مصر، حيث يتشابه بشكل كبير مع لائحة GDPR. ويقوم بتنظيم عمليات جمع وتخزين ومعالجة البيانات الشخصية داخل البلاد.

- متطلبات تطبيقه :
  - الحصول على موافقة الأفراد قبل جمع بياناتهم الشخصية.
  - حماية البيانات الشخصية من التسريب أو الاختراق.
  - تقديم تقارير دورية إلى الهيئة العامة لحماية البيانات.
  - توفير إمكانية الوصول إلى البيانات الشخصية للأفراد والسماح لهم بحذفها إذا طلبوا ذلك.

- **قانون حقوق الملكية الفكرية المصري:** ينظم هذا القانون حماية حقوق الملكية الفكرية للمصنفات العلمية والأدبية إذا كانت البيانات البحثية تحتوي على مواد لها حقوق ملكية فكرية، يجب الالتزام بالقانون المصري لحقوق الملكية الفكرية.

- **متطلبات تنفيذه:**
  - حماية حقوق النشر والتأليف.
  - الحصول على التراخيص اللازمة لاستخدام أي مواد محمية بحقوق الملكية الفكرية.

- **قانون الجريمة الإلكترونية (2018):** ينظم هذا القانون الجرائم المرتبطة بتكنولوجيا المعلومات والاتصالات، ويضع معايير صارمة لحماية البيانات من الوصول غير المصرح به، والجرائم السيبرانية.

- **متطلبات تنفيذه :**
  - حماية مستودعات البيانات من الهجمات السيبرانية.
  - توفير إجراءات أمان تقنية لحماية البيانات المخزنة.

- **قوانين تنظيم الاتصالات:** تقوم هذه القوانين بتنظيم إدارة الاتصالات وتدفق البيانات عبر الإنترنت في مصر. إذا كانت مستودعات البيانات تعتمد على الاتصال بالإنترنت لنقل أو مشاركة المعلومات، فإنه يتوجب الالتزام بقوانين تنظيم الاتصالات.

- **متطلبات تنفيذه:**

- الامتثال للمعايير واللوائح التي تضعها وزارة الاتصالات وتكنولوجيا المعلومات في مصر.
- الحفاظ على سرية البيانات المتداولة عبر الإنترنت.

3/1 - المعايير الأخلاقية والسياسات الجامعية

أولًا: المعايير الأخلاقية للبحث العلمي

- المؤسسات البحثية في مصر عادة ما تضع مدونات سلوك ومعايير أخلاقية لحماية المشاركين في البحث وضمان سرية بياناتهم عند إدارة مستودعات بيانات بحثية، من الضروري الالتزام بهذه المعايير الأخلاقية.

أمثلة على بعض المعايير الأخلاقية المستخدمة في الجامعات المصرية

جدول 1 المعايير الأخلاقية المستخدمة ببعض الجامعات المصرية

جامعة عين شمس	https://science.asu.edu.eg/ar/page/48/أخلاقيات-البحث-العلمي
كلية الزراعة جامعة أسيوط	https://aun.edu.eg/agriculture/ar/postgrad/research

ثانيًا: سياسات الجامعات والمؤسسات البحثية المصرية

تفتقر معظم الجامعات المصرية والمؤسسات البحثية إلى سياسات داخلية تنظم جمع وإدارة البيانات البحثية. من الضروري وضع هذه السياسات وتوثيقها، خاصة عند إدارة مستودع مرتبط بمؤسسة بحثية محلية. ينبغي إعداد خطة إرشادية لإدارة البيانات البحثية، يمكن الاستفادة منها في تطوير الخطط البحثية الخاصة بكل مؤسسة أو مستودع أو مشروع بحثي.

ثالثًا: الاتفاقيات الدولية المتعلقة بالملكية الفكرية

- عند التعامل مع مستودع يحتوي على بيانات بحثية قد تكون محمية بموجب حقوق الملكية الفكرية الدولية، يجب الالتزام باتفاقيات مثل:
- اتفاقية التريبس (TRIPS): تنظم حقوق الملكية الفكرية على المستوى العالمي.
- اتفاقية WIPO: تنظم حقوق المؤلف والتراخيص المتعلقة بالاستخدامات العلمية والأدبية.

رابعًا: التراخيص المفتوحة والبيانات المفتوحة

إذا كانت البيانات البحثية معدة للنشر العام وفقًا لتراخيص البيانات المفتوحة مثل Creative Commons، فمن الضروري الالتزام بشروط هذه التراخيص، والتي تم توضيحها بشكل مفصل في الفصل الخامس من هذه الرسالة. بالإضافة إلى ذلك، ينبغي وضع سياسات

واضحة لإدارة البيانات، تشمل سياسات الخصوصية وحقوق الملكية الفكرية وأمن البيانات، لضمان الامتثال التام للقوانين المحلية والدولية.

خامسًا: السياسات الداخلية

لإنشاء مستودع فعال للبيانات البحثية وإدارته بشكل جيد، من الضروري وضع مجموعة من السياسات الداخلية التي تشمل الجوانب التقنية والقانونية والإدارية. هذا يضمن حماية البيانات، والامتثال للقوانين والمعايير، وتعزيز الاستخدام الأمثل للمستودع. ومن أبرز السياسات الداخلية التي ينبغي وضعها عند إنشاء مستودع للبيانات البحثية:

- سياسة إدارة البيانات (Research Data Management Policy): من الضروري إقرار سياسة وطنية لإدارة البيانات البحثية لتطبيقها في المستودع.

*** عناصر السياسة**

1- سياسة جمع البيانات: يجب أن تتضمن الاستراتيجية سياسة لجمع البيانات ومعاييرها التي سيتم تخزينها في المستودع. ينبغي أن تشمل هذه المعايير نوع البيانات، ومصدرها، وطرق الحصول عليها، بالإضافة إلى الإجراءات اللازمة للحصول على الموافقات لجمع البيانات، خاصة إذا كانت تتعلق بأفراد أو جهات محددة.

2- تنظيم البيانات وتخزينها: يجب تنفيذ إجراءات لتنظيم البيانات وتنسيقها وفقًا للمعايير المعترف بها، مثل معايير FAIR (قابلية البحث، إمكانية الوصول، التشغيل البيني، وإعادة الاستخدام). كما ينبغي تحديد الصيغ المقبولة لتخزين البيانات (مثل CSV، XML، JSON) لضمان تحقيق التشغيل البيني.

3- إجراءات التخزين: يجب وضع سياسات واضحة لتخزين البيانات البحثية على المدى الطويل، مثل تخزين البيانات "الباردة" التي لا تُستخدم بشكل متكرر في وسائط تخزين منخفضة التكلفة، حيث تكون سرعة استرجاع البيانات في هذه الوسائط محدودة يجب وضع سياسات واضحة لتخزين البيانات البحثية على المدى الطويل، مثل تخزين البيانات "الباردة" التي لا تُستخدم بشكل متكرر في وسائط تخزين منخفضة التكلفة، حيث تكون سرعة استرجاع البيانات في هذه الوسائط محدودة (Research Data Management Policy).

كما ينبغي التأكد من إجراء نسخ احتياطي منتظم للبيانات وتوفير خطط لاسترجاعها في حالات الطوارئ (*University of Cambridge Research Data Management Policy Framework*، 2021).

- سياسة الوصول إلى البيانات (Data Access Policy)

أ) مستويات الوصول: تحديد مستويات الوصول إلى البيانات، حيث يتم تصنيف البيانات حسب حساسية المعلومات حيث يمكن أن تكون هناك بيانات متاحة للعامة، وبيانات متاحة فقط للباحثين المعتمدين، وبيانات مقيدة جدًا(" HARVARD RESEARCH DATA SECURITY POLICY", 2012).

ب) آليات طلب الوصول: تحديد آليات واضحة لطلب الوصول إلى البيانات البحثية، بما في ذلك الشروط والإجراءات اللازمة للحصول على تصريح للوصول إلى البيانات الحساسة أو المقيدة. يمكن أن يتم ذلك من خلال تعبئة نماذج خاصة لإتاحة الوصول إلى مواد معينة أو استخدام صلاحيات محددة على المستودع الرقمي. كما يتعين أيضًا استكمال سجل الضيوف الخاص بمجموعات البيانات البحثية قبل تنزيل واستخدام هذه البيانات.

ج) مراقبة الوصول: يجب إنشاء نظام لتسجيل ومراقبة جميع عمليات الوصول إلى البيانات داخل المستودع، بحيث يتضمن سجلًا لكل عملية وصول يوضح من قام بالوصول إلى أي بيانات وتوقيت ذلك. يمكن تحقيق ذلك من خلال أنظمة إدارة المستودعات الخاصة بالبيانات البحثية، باستخدام نظام فرعي مخصص للتقارير وتتبع عمليات الوصول إلى الخادم، أو من خلال الاستعانة ببرمجيات تحليل سجلات الوصول الخاصة بالمستودعات(*Policy on Access to Electronic Information*, 2023).

- سياسة الخصوصية وحماية البيانات (Data Privacy and Security Policy)

أ) حماية البيانات الشخصية: وضع ضوابط صارمة لجمع ومعالجة البيانات الشخصية بما يتوافق مع القوانين المحلية والدولية مثل GDPR وقانون حماية البيانات الشخصية المصري.

- إجراءات تشفير البيانات لحماية البيانات الشخصية والبيانات الحساسة من الوصول غير المصرح به.

ب) إجراءات الأمان السيبراني: وضع معايير للأمان السيبراني، بما في ذلك بروتوكولات حماية الشبكات وأنظمة المستودع من الهجمات الإلكترونية.

- تنفيذ أنظمة كشف التسلل (IDS) وتطبيق جدران حماية قوية لحماية البيانات.

- سياسة حقوق الملكية الفكرية واستخدام البيانات (Intellectual Property and Data Use Policy)

أ) حقوق الملكية الفكرية: تحديد حقوق الملكية الفكرية للبيانات المخزنة في المستودع، بما في ذلك حقوق المؤلفين والباحثين المساهمين في البيانات، إجراءات واضحة لإدارة التراخيص المتعلقة بالبيانات، مثل استخدام التراخيص المفتوحة (Creative Commons) أو التراخيص المقيدة(" Guidance note from the Research Office and Cambridge Enterprise IP Policy (in practice – how it works, who to approach and when?", 2010).

ب) إعادة استخدام البيانات: وضع شروط واضحة لإعادة استخدام البيانات من قبل المستخدمين الخارجيين يمكن أن تشمل هذه الشروط الإشارة إلى المصدر، الحصول على الموافقات اللازمة، واحترام حقوق الملكية الفكرية.

- سياسة الحفظ والتخلص من البيانات (Data Retention and Disposal Policy)

أ) مدة الاحتفاظ بالبيانات: تحديد فترة الاحتفاظ بالبيانات يعتمد على نوعها وأهميتها في البحث. فبعض البيانات قد تتطلب الاحتفاظ بها لفترات طويلة، بينما يمكن التخلص من بيانات أخرى بعد فترة معينة. لذا، من الضروري تحديد مدد حفظ البيانات وفقًا لنوعها واحتياجات استخدامها.(*University of Oxford Policy on the management of research data and records*)

ب) إجراءات التخلص من البيانات: وضع إجراءات واضحة للتخلص الآمن من البيانات بعد انتهاء مدة الاحتفاظ بها، حيث يجب أن تتضمن الاستراتيجية هذه الإجراءات للتأكد من الحذف التام للبيانات من جميع أنظمة التخزين.

- سياسة النسخ الاحتياطي واسترجاع البيانات (Backup and Disaster Recovery Policy)

أ) النسخ الاحتياطي الدوري: إجراءات نسخ احتياطي دورية تضمن حفظ نسخة احتياطية من البيانات بشكل منتظم في مواقع تخزين آمنة.

ب) خطط استرجاع البيانات: إنشاء خطط استرجاع البيانات في حالات الطوارئ (Disaster Recovery Plans)، بما يشمل تحديد الوقت المسموح لاسترجاع البيانات بعد حدوث مشكلة (Recovery Time Objective - RTO).

ج) سياسات الاستجابة للكوارث: وضع خطة استجابة للكوارث لضمان التعامل السريع مع أي اختراق أمني أو فقدان للبيانات، إنشاء فريق للاستجابة للطوارئ يتكون من متخصصين في الأمن السيبراني وإدارة البيانات.

- سياسة الامتثال القانوني والأخلاقي (Legal and Ethical Compliance Policy)

أ) الامتثال للقوانين: الالتزام بجميع القوانين واللوائح المحلية والدولية المتعلقة بإدارة البيانات، مثل قانون حماية البيانات الشخصية المصري وGDPR.

ب) الأخلاقيات البحثية: ضمان الامتثال للمعايير الأخلاقية الخاصة بالبحث العلمي، مثل الحصول على موافقات الأخلاقيات البحثية عند التعامل مع بيانات تتعلق بالأفراد أو الفئات المحمية.

- سياسة التدريب والتوعية (Training and Awareness Policy)

أ) تدريب الموظفين: تنفيذ برامج تدريبية دورية للموظفين المسؤولين عن إدارة المستودع على سياسات حماية البيانات، الأمان السيبراني، والامتثال القانوني.

ب) التوعية العامة: تقديم جلسات توعية للباحثين والمستخدمين حول كيفية استخدام المستودع بشكل آمن وكيفية الالتزام بسياسات الوصول وإدارة البيانات.

- سياسة المراجعة والتقييم (Review and Audit Policy)

أ) المراجعة الدورية: وضع خطة للمراجعة الدورية لسياسات المستودع وعملياته لضمان الامتثال والتحديث حسب الحاجة.

ب) التقييمات الأمنية والتقنية: إجراء تقييمات أمنية وتقنية دورية لتحديد نقاط الضعف وتحديث أنظمة الحماية.

- سياسة الشراكات والتعاون (Partnership and Collaboration Policy)

أ) إدارة الشراكات: وضع إطار قانوني للشراكات مع المؤسسات البحثية الأخرى، لضمان تبادل البيانات بشكل قانوني وآمن.

ب) التكامل مع مستودعات خارجية: ضمان توافق المستودع مع المعايير الدولية لتبادل البيانات البحثية يسهم في تسهيل التكامل مع مستودعات البيانات الأخرى على المستويات المحلية والدولية. إن وضع هذه السياسات يعزز من إدارة فعالة ومستدامة لمستودع البيانات

البحثية، مع الالتزام بأفضل الممارسات والمعايير الدولية لحماية البيانات، مما يضمن سهولة الوصول إليها وتشغيلها بشكل آمن وفعال.

4/1 البنية التحتية والتكنولوجية

عند إنشاء مستودع خاص بالبيانات البحثية، يجب أخذ البنية التحتية بعين الاعتبار كعنصر أساسي لنجاح واستدامة المشروع. تُعتبر الأنظمة والمستودعات الرقمية مفتوحة المصدر من الأدوات الحيوية التي تدعم حركة العلوم المفتوحة وتعزز إدارة البيانات البحثية. حيث توفر هذه الأنظمة مرونة كبيرة وإمكانية تعديل وتطوير البرمجيات لتلبية احتياجات المجتمع الأكاديمي والبحثي دون قيود من شركات البرمجيات. كما تسهم هذه الأنظمة في تعزيز الشفافية وتسهيل التعاون بين الباحثين من خلال مشاركة الأدوات والموارد البحثية بشكل مفتوح، مما يزيد من فرص تكرار النتائج والتحقق منها (Naheem & Mir, 2024).

تُعتبر المستودعات الرقمية منصة أساسية لحفظ وتنظيم ونشر البيانات البحثية، حيث توفر بنية تحتية تضمن الوصول المفتوح إلى هذه البيانات على المدى الطويل، مما يسهم في إعادة استخدامها وتوظيفها في الأبحاث المستقبلية (Tennant et al., 2016).

- **السمات العامة للبنية التحتية:** يجب وضع النقاط الرئيسية التالية في الاعتبار عند تهيئة البنية التحتية للمستودع الرقمي للبيانات البحثية

أولا قابلية التوسع: الاستثمار في بنية تحتية تكنولوجية قابلة للتوسع لمعالجة كم كبير من البيانات، مع ضمان الموثوقية والأداء والأمان مع وضع الحلول السحابية في الاعتبار لتعزيز المرونة والقابلية للتوسع.

ثانيا التكامل: ضمان أن المستودع يمكن أن يتكامل مع الأنظمة الوطنية والدولية الأخرى، بما في ذلك المستودعات والشبكات الأخرى واعتماد بروتوكولات ومعايير البيانات القياسية، مثل OAI-PMH (بروتوكول أرشيف المبادرات المفتوحة لجمع البيانات الوصفية) وRDF (إطار وصف الموارد).

- **المعيارية:** تضمن معيارية البنية التحتية جودة التشغيل والقدرة على قياس الأداء وضمان الاستمرارية لمستودع البيانات البحثية، بالإضافة الي ضبط كيفية تشغيل المستودع وضمان كفاءة العمليات المختلفة وضمان الاستدامة طويلة الأجل(Boyd, 2021) .

- **التجهيزات الحاسوبية:** يفضل استخدام الخوادم السحابية نظرًا لما تقدمه من استقرار وإتاحة وأمان أعلى مقارنة بالخوادم المحلية، بالإضافة إلى تقليل الحاجة إلى الكوادر البشرية لإدارة وتأمين هذه الخوادم.
- **الخوادم السحابية:** يمكن الاستفادة من خدمات الحوسبة السحابية مثل Amazon Web Services (AWS) وMicrosoft Azure وGoogle Cloud لتحقيق التوسع والمرونة. يمكن استضافة Dataverse على خدمات مثل EC2 في Amazon أو VM في Azure، مع إمكانية التكامل مع خدمات التخزين السحابية.
- **التخزين السحابي**: تخزين البيانات على حلول سحابية مثل Amazon S3 أو Google Cloud Storage لضمان التخزين الآمن والموثوق للبيانات البحثية

1/ 5 النموذج والمواصفات الموصى بها لنظام إدارة المستودعات الرقمية Dataverse

تعتبر المواصفات المعتمدة في جامعة هارفارد مرجعًا يمكن للجامعات والمؤسسات الاستناد إليه عند تشغيل نظام Dataverse. تُعد Dataverse منصة مرنة لإدارة البيانات البحثية، يمكن تشغيلها على أجهزة ذات مواصفات بسيطة. على سبيل المثال، يمكن تشغيل نسخة تجريبية من النظام على جهاز افتراضي مزود بمعالجين بسرعة 2.8 جيجاهرتز وذاكرة وصول عشوائي بسعة 8 جيجابايت و50 جيجابايت من مساحة التخزين. ومع ذلك، يُفضل استخدام تجهيزات أكثر قوة لضمان الأداء الأمثل في بيئات الإنتاج في حالة تشغيل النظام على خوادم محليه فإن المواصفات الملائمة هي مجموعة خوادم تحتوي على معالجات Intel Xeon بسرعة 2.53 تتكون من ثلاث خوادم بذاكرة 48 جيجابايت تُستخدم كواجهات ويب تعتمد على Glassfish 4 وApache، مع توزيع الحمل بواسطة خادم توزيع الحمل. وثلاث خوادم أخرى مخصصة لقواعد البيانات وتطبيق Rserve، مع تخزين متعدد التيرابايت يتم عبر نظام تخزين شبكي (SAN) باستخدام بروتوكول.

يمكن استخدام الخوادم السحابية من شركة أمازونAWS كبديل للخوادم المحلية ببنية تتكون من أربعة خوادم افتراضية من خلال Amazon Web Services (AWS). حيث تُشغل واجهات الويب على جهازين من نوع "m4.4xlarge" بذاكرة 64 جيجابايت و16 وحدة معالجة مركزية افتراضية (vCPUs)، مع توزيع الحمل عبر موازن تحميل AWS ELB. بينما يُشغل محرك

البحث Solr على جهاز من نوع "m4.2xlarge" بذاكرة 32 جيجابايت و8 وحدات معالجة افتراضية. يعتمد تطبيق R على جهاز "m4.xlarge" بذاكرة 16 جيجابايت و4 وحدات معالجة.

أما قاعدة البيانات PostgreSQL فتعمل عبر خدمة Amazon RDS، في حين تُخزن الملفات على Amazon S3 مع تكرار فوري للبيانات في وسائط تخزين ثانوية ودعم النسخ الاحتياطي على Glacier، مما يسهل استرجاع البيانات في حالة الحذف. يُنصح بتجربة تكوينات مختلفة للأجهزة بما يتناسب مع حجم البيانات واحتياجات المستخدم وهذه هي البنية المعتمدة في جامعة هارفارد (King, 2007).

الأنظمة والتجهيزات البرمجية

- **أنظمة إدارة المستودعات الرقمية:** تعتبر مستودعات إدارة البيانات البحثية الأداة الأساسية لإدارة البيانات في مجال العلوم المفتوحة، حيث توفر بنية تحتية فعالة لحفظ وتنظيم ونشر البيانات البحثية بطرق تسهل إعادة استخدامها وتعزز من نطاق الاستفادة منها. تتيح هذه المستودعات للباحثين أرشفة بياناتهم ونشرها بشكل موثوق وآمن، مع إمكانية إنشاء معرفات دائمة (DOIs) تسهل الإشارة إليها في الأبحاث المستقبلية. من أبرز أمثلة مستودعات البيانات البحثية: Dataverse وDryad وZenodo، حيث يساهم كل منها في حفظ وتوزيع البيانات بطرق مفتوحة وآمنة.

- **مستودع Dataverse:** تُعتبر هذه البرمجية أداة لإدارة البيانات البحثية، تم تطويرها في معهد العلوم الكمية الاجتماعية بجامعة هارفارد. تتيح للباحثين تحميل وتنظيم مجموعات بياناتهم مع إعدادات للتحكم في الوصول، مما يمكّنهم من تقديم البيانات بشكل مفتوح أو مقيد حسب الحاجة (King, 2007). بالإضافة إلى ذلك، يوفر المستودع إمكانية الاستشهاد بالبيانات باستخدام معرفات دائمة، مما يعزز الشفافية ويزيد من استدامة الأبحاث العلمية. يسهم هذا النوع من المستودعات في تحسين التعاون البحثي من خلال تمكين الباحثين من مشاركة بياناتهم والاستفادة من البيانات المنشورة من قبل الآخرين. كما يعزز موثوقية البحث العلمي من خلال تسهيل التحقق من النتائج وإعادة استخدامها (Piwowar & Vision, 2013).

- **البرمجيات المساعدة :** هناك العديد من الأدوات الخارجية التي تتكامل مع مستودعات إدارة البيانات البحثية، سواء في مرحلة التخطيط أو الإنشاء، أو في المراحل السابقة لوضع البيانات البحثية في المستودعات. تشمل هذه الأدوات برمجيات تجهيز وجمع البيانات، بالإضافة إلى

البرمجيات المساعدة التي يمكن استخدامها لاحقًا كإضافات لعرض أنواع معينة من البيانات أو أدوات تحليل البيانات. وقد تم استعراض هذه الأدوات بشكل كامل في الفصل الثالث من الرسالة.

1/ 6 الصيانة والتحديث

تعتبر الصيانة والدعم الفني الأساس الذي يقوم عليه تطبيق الأنظمة والتطبيقات المستخدمة في إدارة البيانات البحثية. وبخاصة إذا كانت هذه التطبيقات مفتوحة المصدر، فإن الأمر يتطلب تنفيذ مهام منتظمة ومحددة لتقديم خدمات الدعم الفني والصيانة والتحديث. تشمل عمليات الدعم الفني إصلاح الأخطاء في الأنظمة، وتحديثها، ومعالجة الثغرات الأمنية، بالإضافة إلى تقديم الدعم اليومي والاستشارات المتعلقة بتطبيق النظام. كما يتعين متابعة عمليات النسخ الاحتياطي والتأكد من تنفيذها بشكل صحيح لضمان حماية النظام في حال حدوث كوارث تهدد البيانات والأنظمة. من الضروري وضع متطلبات عقد الصيانة والدعم الفني مع مواصفات تنفيذ النظام قبل بدء المشروع، وذلك لضمان استدامته وتوفير الموارد المالية اللازمة لاستمراره. بالإضافة إلى ذلك، يجب تحديد التفاصيل ومتطلبات المؤسسة للخدمات المقدمة خلال فترة الدعم الفني، مثل التطوير، والتكامل مع الأنظمة الأخرى، والتكامل مع أنظمة التوثيق والدخول الموحد، وإعداد التقارير، ومتابعة مصادر الخوادم، وغيرها من الخدمات الضرورية لضمان استمرارية عمل الأنظمة بكفاءة وفعالية.

يمكن أيضًا إدراج بعض البرامج التدريبية التي تهدف إلى تعليم كيفية استخدام المستودع الرقمي وإدارته، بالإضافة إلى التدريب المستمر على الوظائف الجديدة والميزات التي تُضاف مع التحديثات الجديدة للأنظمة (Anbalagan, 2024).

7/1 سياسات تهيئة وضبط بيانات المستودع وإدارته

وتعد سياسات تهيئة المستودعات الخاصة بالبيانات البحثية واحدة من أهم عناصر إنشاء مستودع بيانات بحثية قوي وفعال ومن أهم هذه السياسات

- تقسيم البيانات وتحديد الموضوعات الرئيسية والفرعية للمستودع.
- تحديد مستويات سرية البيانات (مواد للنشر العام، مواد منخفضة السرية، مواد متوسطة السرية، مواد عالية السرية).
- تحديد قوالب مجموعات البيانات وتوصيفها.
- تحديد بيانات قوالب دفتر الزوار Gustbook.
- تحديد أنواع عناصر البيانات الوصفية للبيانات البحثية .
- تحديد تنسيقات البيانات المستخدمة لملفات البيانات البحثية (pdf, word , rdf, json txt, excel,) وغيرها من أنواع الملفات.
- تحديد قوالب البيانات الوصفية والمعايير المستخدمة فيها مثل DDI، TEI، EML، MARC، CMDI.
- تحديد الأدوار والصلاحيات الخاصة بمجموعات البيانات أو الملفات أو المجموعات الفرعية.
- تحديد الأدوات الإضافية للتعامل مع البيانات البحثية مثل رفع وعرض الملفات وتنظيم مسارات العمل وأدوات الاستعلام والتحليل.
- تنظيم الملفات الخاصة بالبيانات البحثية ووضع نظام لتسمية الملفات Naming convention
- تحديد نظام وآلية إدارة الإصدارات الخاصة بالبيانات البحثية.
- تحديد APIsالمستخدمة في التكامل مع موقع المؤسسات أو الهيئات المستخدمة للمستودع أو التكامل مع أي منصات أخرى.
- اختيار وتحديد المعرفات الدائمة المستخدمة في المستودعات وآلية التكامل مع منصاتها سواء كانت للبيانات البحثية أو للباحثين.

يتضمن الفصل الخامس من الرسالة أمثلة متكاملة على السياسات الخاصة بمستودعات البيانات البحثية.

معايير البيانات والبيانات الوصفية:

معايير البيانات الوصفية: يجب تطبيق المعايير الدولية المعترف بها للبيانات الوصفية (مثل DataCite وDublin Core وISO 19115) لضمان توفير وصف شامل ومتسق للبيانات.

كما أن اختيار وتحديث البيانات الوصفية لمجموعات وعناصر البيانات في مستودعات البيانات البحثية يعد عملية أساسية لضمان تنظيم وتوثيق البيانات بشكل صحيح وفعال. فهذه البيانات الوصفية تلعب دورًا مهمًا في تصنيف البيانات، وتسهيل الوصول إليها، وتعزيز إمكانية إعادة استخدامها (Asok & Dandpat, 2023).

الوصول والحفظ:

- سياسات الوصول: يجب وضع سياسات وصول واضحة توازن بين الانفتاح والخصوصية والأمان، مع ضرورة تنفيذ آليات للتحكم في الوصول ومشاركة البيانات. ينبغي أن تتماشى هذه السياسات مع اتفاقيات مشاركة البيانات الدولية، وأن تميل نحو تطبيق الممارسات العالمية المتعلقة بالعلوم المفتوحة، دون أن تتعارض مع الخصوصية والسرية المطلوبة للبحوث ذات الطبيعة الحساسة. كما يجب دعم خصائص النشر، مثل تأخير الإتاحة (embargo) وغيرها من الممارسات التي تضمن حقوق الباحثين في إطار مبادرات العلوم المفتوحة.

- الحفظ طويل الأمد: يتطلب الأمر وضع استراتيجيات للحفاظ على البيانات على المدى الطويل، بما في ذلك النسخ الاحتياطي المنتظم والاحتفاظ بالنسخ الاحتياطية. يجب الالتزام بالمعايير الدولية للحفظ وأفضل الممارسات، مثل نموذج OAIS (نظام المعلومات الأرشيفية المفتوحة) (Sofi et al., 2024).

التعاون والمشاركة:

يجب على مستودعات البيانات البحثية أن تتفاعل مع الأطراف المعنية في مجال البحث العلمي على المستويين الوطني والدولي، بما في ذلك المجتمعات البحثية ووكالات التمويل وصناع السياسات.

بالإضافة إلى ذلك، ينبغي تعزيز التعاون من خلال إقامة شراكات والمشاركة في المبادرات والشبكات الدولية.

مشاركة المجتمع: من الضروري أيضًا إشراك المجتمع البحثي في عملية تطوير وتحسين المستودع بشكل مستمر، من خلال جمع الملاحظات والتكيف مع الاحتياجات والتطورات التكنولوجية الجديدة.

1/ 8 تقييم الأداء

- **مؤشرات الأداء:** أصبح تقييم الأداء عنصراً أساسياً لضمان نجاح واستمرارية المشاريع، خاصةً إذا كنا نهدف إلى إنشاء مستودع خاص بإدارة البيانات البحثية يتسم بالاستدامة وجودة الخدمة المقدمة. لذا، من الضروري وضع معايير لتقييم الأداء الخاص بالمستودعات أثناء عملية إنشائها. يعتبر تطوير مقاييس ومؤشرات أداء رئيسية (KPIs) لتقييم فعالية وتأثير المستودع، والتي تشمل إحصاءات الاستخدام، ورضا المستخدمين، وقياسات جودة البيانات، من الأسس الضرورية لتقييم المستودعات. لتقييم فعالية وتأثير مستودع البيانات البحثية، يمكن الاعتماد على مجموعة من مؤشرات الأداء الرئيسية (KPIs) التي تغطي جوانب متعددة من الأداء. يساعد نموذج KPIs في قياس النجاح، وتحليل الأداء، وتحسين العمليات. فيما يلي نموذج مفصل لمؤشرات الأداء الرئيسية لتقييم مستودع البيانات البحثية (Ball & Duke, 2015).

أولًا: مؤشرات قياس الأداء المتعلقة بالاستخدام

أ. عدد المستخدمين النشطين

- التعريف: عدد المستخدمين الذين قاموا بتسجيل الدخول إلى المستودع أو استخدموا البيانات خلال فترة زمنية معينة (شهريًا، سنويًا).
- الهدف: زيادة عدد المستخدمين النشطين بمرور الوقت.

ب. عدد عمليات البحث والاستعلامات

- التعريف: إجمالي عدد عمليات البحث والاستعلامات التي يتم تنفيذها في المستودع.
- الهدف: زيادة عدد عمليات البحث لتحسين قابلية اكتشاف البيانات.

ج. عدد عمليات تحميل البيانات

- التعريف: عدد مرات تحميل البيانات من المستودع.

- الهدف: زيادة عدد عمليات التحميل لقياس مدى اهتمام المستخدمين بالبيانات المتاحة.

ثانيًا: مؤشرات قياس الأداء المتعلقة بجودة البيانات

أ. دقة البيانات الوصفية

- التعريف: نسبة العناصر الوصفية التي تتوافق مع المعايير والمواصفات المحددة.
- الهدف: تحقيق نسبة عالية من الدقة في البيانات الوصفية.

ب. مستوى استكمال البيانات الوصفية

- التعريف: نسبة الحقول الوصفية المكتملة مقابل الحقول التي يتم تركها فارغة.
- الهدف: زيادة نسبة الحقول المكتملة لضمان توفير معلومات شاملة.

ج. عدد الأخطاء والتعديلات في البيانات الوصفية

- التعريف: عدد الأخطاء والتعديلات التي تم اكتشافها وإصلاحها في البيانات الوصفية.
- الهدف: تقليل عدد الأخطاء وتحسين جودة البيانات.

ثالثًا: مؤشرات قياس الأداء المتعلقة بالأداء الفني

أ. وقت تحميل الصفحات واستجابة النظام

- التعريف: الوقت الذي يستغرقه تحميل الصفحات أو استجابة النظام لطلبات المستخدم.
- الهدف: تحسين وقت تحميل الصفحات واستجابة النظام لضمان تجربة مستخدم سلسة.

ب. توافر النظام ووقت التعطل

- التعريف: نسبة الوقت الذي يكون فيه النظام متاحًا للمستخدمين مقابل الوقت الذي يكون فيه النظام غير متاح بسبب أعطال أو صيانة.
- الهدف: الحفاظ على نسبة توافر عالية للنظام.

رابعًا: مؤشرات قياس الأداء المتعلقة بالبحث والتصفح

أ. معدل نجاح البحث

- التعريف: نسبة البحث الذي يؤدي إلى استرجاع بيانات ذات صلة وملائمة مقارنة بالبحث الذي لا يؤدي إلى نتائج مفيدة.
- الهدف: زيادة معدل نجاح البحث لتحسين فعالية أدوات البحث.

ب. عدد البيانات المكتشفة

- التعريف: عدد العناصر أو مجموعات البيانات التي تم اكتشافها من خلال البحث.
- الهدف: زيادة عدد البيانات المكتشفة لتحسين إمكانية اكتشاف البيانات.

خامسًا: مؤشرات قياس الأداء المتعلقة بتأثير البيانات

أ. عدد الاقتباسات أو الاستخدامات البحثية

- التعريف: عدد المرات التي يتم فيها اقتباس البيانات أو استخدامها في الأبحاث المنشورة.
- الهدف: زيادة عدد الاقتباسات أو الاستخدامات لتقييم التأثير الأكاديمي والبحثي.

ب. ردود الفعل من المستخدمين

- التعريف: جمع وتحليل ملاحظات المستخدمين حول جودة وملاءمة البيانات والخدمات المقدمة.
- الهدف: تحسين تجربة المستخدم بناءً على الملاحظات والاقتراحات.

سادسًا: مؤشرات قياس الأداء المتعلقة بالأمان والامتثال

أ. عدد حوادث الأمان

- التعريف: عدد الحوادث الأمنية مثل خروقات البيانات أو محاولات الوصول غير المصرح بها.
- الهدف: تقليل عدد حوادث الأمان لضمان حماية البيانات.

ب. الامتثال للمعاييروالسياسات

- التعريف: نسبة امتثال النظام للمعايير والسياسات المحلية والدولية.
- الهدف: تحقيق نسبة امتثال عالية لضمان التوافق مع السياسات القانونية والأخلاقية(Mosha & Ngulube, 2023).

سابعًا: مؤشرات قياس الأداء المتعلقة بالصيانة والتحديثات

أ. عدد التحديثات والتعديلات

- التعريف: عدد التحديثات والتعديلات التي تم إجراؤها على النظام والبيانات.
- الهدف: تحسين النظام بانتظام من خلال تحديثات وتعديلات فعالة.

ب. وقت استجابة الدعم الفني

- التعريف: الوقت الذي يستغرقه فريق الدعم الفني لحل المشكلات التي يواجهها المستخدمون.
- الهدف: تقليل وقت استجابة الدعم الفني لضمان معالجة المشكلات بسرعة(Aithal & Aithal, 2023).

ثامنًا: التمويل والاستدامة

نموذج التمويل: يتضمن تأمين التمويل من خلال الدعم الحكومي، والمنح، والشراكات، بالإضافة إلى تطوير نموذج مالي مستدام يضمن التشغيل والصيانة على المدى الطويل للمستودع. لإنشاء نموذج تمويل مستدام لإدارة البيانات البحثية وتأسيس مستودع وطني للبيانات، يجب تحديد مصادر التمويل وتنويعها، بما في ذلك الجهات المانحة والمؤسسات الخارجية التي قد تكون مهتمة بإجراء بحوث ودراسات وقادرة على تمويلها، فضلاً عن تمويل المستودعات المخصصة لتخزين البيانات المتعلقة بهذه المشاريع، بالإضافة إلى الدعم الحكومي المخصص للبحث العلمي.

تاسعًا: الرؤية والأهداف

يجب تحديد رؤية واضحة للمستودع الوطني للبيانات البحثية، مثل إنشاء مستودع وطني يهدف إلى إدارة البيانات البحثية بشكل يضمن الوصول المستدام والآمن إليها، مع تعزيز التعاون العلمي بين الباحثين والمؤسسات البحثية.

يجب أيضا صياغة الأهداف الخاصة بالمستودع لتشمل:

- دعم الحفاظ على البيانات البحثية وإدارتها بشكل فعال.
- تحسين تبادل البيانات وتعزيز الابتكار.
- ضمان تمويل مستدام يضمن استمرارية المشروع على المدى الطويل.

عاشرًا: التخطيط المالي

- مصادر التمويل المحتملة

*** التمويل الحكومي:**

- تقديم مقترحات إلى الحكومة الوطنية والهيئات البحثية للحصول على دعم مباشر.
- الاستفادة من خطط تمويل البحث والتطوير التي تقدمها الوزارات أو هيئات العلوم والتكنولوجيا.

*** الشراكات مع القطاع الخاص:**

- التعاون مع الشركات التقنية أو المؤسسات التي تهتم بتوفير حلول إدارة البيانات مثل شركات البرمجيات والتكنولوجيا.
- نموذج الشراكة العام-الخاص (PPP) لتأمين جزء من التمويل مقابل حصول الشركات على بيانات البحث المصرح بها.

*** التمويل الدولي:**

- التقدم للحصول على منح من منظمات دولية مثل الاتحاد الأوروبي، البنك الدولي، أو اليونسكو لدعم مشاريع إدارة البيانات المفتوحة والبحثية.
- استكشاف برامج تمويل مثل Horizon Europe أو NSF.

*** الاشتراكات والعضويات:**

- تقديم نماذج اشتراك مدفوعة للباحثين والمؤسسات الأكاديمية للوصول إلى خدمات المستودع، مع الحفاظ على نموذج مفتوح للمشاركة العامة في جوانب معينة.

*** التبرعات والمنح:**

- إطلاق حملات تبرع ودعوة المؤسسات الأكاديمية والبحثية الكبرى للتبرع لدعم المشروع.
- إنشاء صندوق دعم مجتمعي لدعم النمو المستدام.

*** التكاليف التقديرية**

- التكاليف الأولية: تتضمن تطوير البنية التحتية التقنية، وشراء أو تطوير البرمجيات، وتوظيف الموظفين الفنيين والإداريين.
- التكاليف التشغيلية: تشمل صيانة الأنظمة، الأمن السيبراني، تكاليف تخزين البيانات على المدى الطويل، وتحديث البرامج.
- التكاليف الإدارية: الرواتب، تكاليف التدريب والتطوير المهني للفريق، والنفقات العامة مثل الكهرباء والاتصالات.

*** الهيكل التنظيمي:**

أ) إدارة المشروع

- اللجنة التنفيذية: تتكون من ممثلين عن الجامعات الكبرى، الهيئات البحثية، والمؤسسات الحكومية.

- فريق التكنولوجيا: مسؤول عن تطوير النظام وتحديثه وتقديم الدعم الفني.

- فريق إدارة البيانات: مسؤول عن السياسات الخاصة بإدارة البيانات وضمان الالتزام بالمعايير الوطنية والدولية.

- فريق التمويل والشراكات: يعمل على تأمين التمويل المستدام من مصادر مختلفة.

ب) استراتيجية الحوكمة

- إنشاء إطار قانوني واضح لضمان الملكية الفكرية وحماية الخصوصية.

- اعتماد سياسات واضحة لإدارة البيانات بما يتوافق مع المعايير الدولية مثل FAIR (Findable, Accessible, Interoperable, Reusable).

ج) استراتيجية الاستدامة المالية

- تنويع مصادر الدخل: إنشاء منصة متكاملة تقدم خدمات متعددة مثل تقديم تقارير تحليلية، دعم البحث العلمي، والتدريب على إدارة البيانات. توفير خيارات إضافية مدفوعة مثل استشارات إدارة البيانات أو تخصيص مساحات تخزين إضافية للمؤسسات.

- استراتيجيات خفض التكاليف: استغلال خدمات الحوسبة السحابية لتقليل تكاليف الصيانة الداخلية، التعاون مع المؤسسات التعليمية والبحثية لتقاسم الموارد والبنية التحتية.

خاتمة

ناقش البحث القوانين والاتفاقيات المنظمة لإنشاء مستودع للبيانات البحثية ، سواء على المستوى الدولي أو المحلي، بإضافة إلى المعايير الأخلاقية والسياسات الجامعية التي ستنظم عمليات إدارة البيانات البحثية ومشاركتها. كما تناول بالتفصيل سياسات إدارة البيانات، بدءًا من جمع البيانات وتنظيمها وتخزينها، وصولًا إلى إتاحتها، ومراقبة الوصول إليها، وتأمينها، والنسخ الاحتياطي، وغيرها من الإجراءات اللازمة لضمان الحماية والحوكمة الفعالة للبيانات.

تطرق البحث أيضًا إلى البنية التحتية والتكنولوجية المطلوبة لإنشاء المستودع الوطني للبيانات البحثية، مسترشدًا بالبنية التحتية الخاصة بجامعة هارفارد في تطبيق منظومة داتافيرس .(Dataverse) كما استعرض الأنظمة والتجهيزات البرمجية الخاصة بإدارة البيانات البحثية، بما في ذلك أنظمة المستودعات والبرمجيات المساندة، واحتياجاتها التشغيلية لضمان استمرارية عملها بكفاءة. إضافة إلى ذلك، ناقش البحث سياسات ضبط المستودع وتهيئته، ومعايير البيانات الوصفية المستخدمة، ثم تناول تقييم أداء المستودع من خلال وضع مؤشرات لقياس فعاليته وكفاءته.

وفي الختام، قدم البحث نموذجًا تطبيقيًا لإدارة المستودع الوطني للبيانات البحثية، يشمل التخطيط المالي، والتشغيلي، والإداري لضمان كفاءة عمل المستودع واستدامته، بما يحقق الأهداف المرجوة في دعم البحث العلمي وتعزيز مشاركة البيانات البحثية على المستويين المحلي والدولي.

مصادر البحث

قانون حماية حقوق الملكية الفكرية رقم 82 لسنة 2002, (٢٠٢٠).

About SciNote - scinote. Retrieved 08-05-2023 from https://scinote.readthedocs.io/en/latest/About-SciNote/

Adobe Bridge. Retrieved 08-05-2023 from https://www.adobe.com/eg_en/products/bridge.html

Aithal, S., & Aithal, S. (2023). Key Performance Indicators (KPI) for Researchers at Different Levels & Strategies to Achieve it. *International Journal of Management, Technology, and Social Sciences*, 294-325. https://doi.org/10.47992/IJMTS.2581.6012.0304

Alsufaian, R., Almoussa, R., Alghamdi, R., Alqahtani, K & ,.Alajmi, R. (2022). *Secure File Storage On Cloud Using Hybrid Cryptography* .

analysis-ready-datasets. https://datamanagement.hms.harvard.edu/analyze/analysis-ready-datasets

Ball, A. (2014). *How to License Research Data*. Digital Curation Centre .

Bhandari, P. (2022). *What Is Data Cleansing? | Definition, Guide & Examples*. Retrieved 02-05-2023 from https://www.scribbr.com/methodology/data-cleansing/

Boyd, C. (2021). Understanding Research Data Repositories as Infrastructures. *Proceedings of the Association for Information Science and Technology*, *58*, 25-35. https://doi.org/10.1002/pra2.433

Bulk Rename Utility. Retrieved 08-05-2023 from https://www.bulkrenameutility.co.uk/#features

Carole, & Goble. FAIR

Computational Workflows .

Chen, H.-l., Chiu, T.-H & ,.Cline, E. (2022). Academic libraries and research data management: a case study of Dataverse global adoption. https://doi.org/10.1108/idd-04-2022-0028

dans. Retrieved 3-11-2024 from https://dans.knaw.nl/en/

Data Analysis in Research: Types & Methods | QuestionPro. In.

Data Cleaning Techniques in Data Mining and Machine Learning. Retrieved 02-05-2022 from https://www.projectpro.io/article/data-cleaning-techniques/651

data-security. In: Harvard university.

Dataverse Project. (2025). https//:sphinx.dataverse.org/about/index.html

Dataverse user guide. Retrieved 27-04-2024 from https://guides.dataverse.org/en/5.13/user/account.html#id13

Dataverse website. Retrieved 28-04-2024 from https://dataverse.org

DDI : learn. https://ddialliance.org/learn

DMP tool. https://dmptool.org

Donaldson, D. R., & Koepke, J. W. (2022). A focus groups study on data sharing and research data management. In (Vol. 9.(

Doty, J., Kowalski, M., & Nash, B. (2015). Making Student Research Data Discoverable: A Pilot Program Using Dataverse. *Journal of Librarianship and Scholarly Communication*, *3*. https://doi.org/10.7710/2162-3309.1234

EUROPE, S. (2021). *Practical Guide to the International Alignment of Research Data Management' (Extended Edition)*. SCIENCE EUROPE .

Foster, E. D. (2017). Open Science Framework (OSF). *Journal of the Medical Library Association*(105(2)). https://doi.org/dx.doi.org/10.5195/jmla.2017.88

Frąckiewicz, M. (2023). *Cloud Storage and Archiving for Scientific Research and Data Analysis*. TS2 SPACE .Retrieved 02-05-2023 from

Gajbe, S. B., Tiwari, A., Gopalji, & Singh, R. K. (2021). Evaluation and analysis of Data Management Plan tools: A parametric approach. *Information Processing & Management*, *58*(3), 102480. https://doi.org/https://doi.org/10.1016/j.ipm.2020.102480

Guidance note from the Research Office and Cambridge Enterprise IP Policy in practice – how it works, who to approach and when? (2010).

https://www.enterprise.cam.ac.uk/wp-content/uploads/2015/04/ip-policy-in-practice-guidance-note-25may10.pdf

Guide To Data Cleaning: Definition, Benefits, Components, And How To Clean Your Data. Retrieved 2023 from https://www.tableau.com/learn/articles/what-is-data-cleaning

Guidelines for Responsible and Sustainable Research Data Management at KIT. (2 .(016 Karlsruher. https://www.rdm.kit.edu/downloads/KIT-RDM-Policy.pdf

Harvard Information security and data privacy. Retrieved 27-7-2024 from https://privsec.harvard.edu/data-classification-table

HARVARD RESEARCH DATA SECURITY POLICY. (2012). https://files.vpr.harvard.edu/files/vpr-documents/files/hrdsp_10_14_14_final_edits.pdf

Image Management | Data Management. In.

Inau, E. T., Sack, J., Waltemath, D., & Zeleke, A. A. (2021). Initiatives, Concepts, and Implementation Practices of FAIR (Findable, Accessible, Interoperable, and Reusable) Data Principles in Health Data Stewardship Practice: Protocol for a Scoping Review. *JMIR Research Protocols*, *10*(2), e22505. https://doi.org/10.2196/22505

Kanza, S., Willoughby, C., Gibbins, N., Whitby, R., Frey, J ,.Erjavec, J.,...Kovac, K. (2017). Electronic Lab Notebooks: can they replace paper? *Journal of Cheminformatics*, *9*. https://doi.org/10.1186/s13321-017-0221-3

Luprich, T. (2020). *Octoparse : Web ScrapOctoparse : Web Scraping Handbooking Handbook* Octoparse .

Mendely blog. Retrieved 08-05-2023 from https://blog.mendeley.com/

Mosha, N., & Ngulube, P. (2023). The utilisation of open research data repositories for storing and sharing research data in higher learning institutions in Tanzania. *Library Management*, *4* .4https://doi.org/10.1108/LM-05-2023-0042

Mushi, G. E., Pienaar, H., & van Deventer, M. (2020). Identifying and Implementing Relevant Research Data Management Services for the Library at the University of Dodoma, Tanzania. *Data Science Journal*, *19*. https://doi.org/10.5334/dsj-2020-001

openrefine.org. https://openrefine.org/docs

Organising your data. In.

Policy framework on research data. Retrieved 26-7-2024 from https://www.ukri.org/who-we-are/epsrc/our-policies-and-standards/policy-framework-on-research-data/

Policy on Access to Electronic Information. (2023). Harvard College .

Policy paper National Data Strategy Mission 1 Policy Framework: Unlocking the value of data across the economy. (2021). Retrieved 3-11-2024 from https://www.gov.uk/government/publications/national-data-strategy-mission-1-policy-framework-unlocking-the-value-of-data-across-the-economy/national-data-strategy-mission-1-policy-framework-unlocking-the-value-of-data-across-the-economy

Ram, K. (2013). Git can facilitate greater reproducibility and increased transparency in science. *Source Code for Biology and Medicine*, *8*(1), 7. https://doi.org/10.1186/1751-0473-8-7

rdm organiser. Retrieved 08-05-2023 from https://rdmorganiser.github.io

Research Data Ownership Policy. (2019). In :H ARVARD U NIVERSITY.

Research data Oxford. Retrieved 28-7-2024 from https://researchdata.ox.ac.uk

Research Data Service University of Edinburgh. Retrieved 28-7-2024 from https://www.ed.ac.uk/information-services/research-support/research-data-service

Research Software Directory. (2025). Retrieved 3-3-2025 from https://research-software-directory.org/

researchspace.com. Retrieved 08-05-2023 from https://www.researchspace.com/

Riley, J. (2017). *UNDERSTANDING METADATA WHAT IS METADATA, AND WHAT IS IT FOR?*National	Information	Standards	Organization .

Share & Publish. Harvard University. Retrieved 5-2-2023 from https://datamanagement.hms.harvard.edu/share-publish

Share and Manage Your Research Data. (2019). [Article]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(12), 5404-5404. https://doi.org/10.1109/jstars.2020.2970774

Snyder, J. (2019). Data Cleansing: An Omission from Data Analytics Coursework. *Information Systems Education Journal*, *17*, 22-29 .

Storage During Research. https://libguides.vu.nl/rdm/data-storage

Tedersoo, L., Küngas, R., Oras, E., Köster, K., Eenmaa, H., Leijen, Ä.,...Sepp, T. (2021). Data sharing practices and data availability upon request differ across scientific disciplines. *Scientific Data*, *8*(1). https://doi.org/10.1038/s41597-021-00981-0

Thorne, S. (2000). Data analysis in qualitative research. *Evidence-Based Nursing*, *3*(3), 68-70. https://doi.org/10.1136/EBN.3.3.68

Top Referencing Tools and Reference Management Software for Academic Writing - Ref-n-Write: Scientific Research Paper Writing Software Tool - Improve Academic English Writing Skills. In.

Trisovic, A., Durbin, P., Schlatter, T., Durand, G., Barbosa, S., Brooke, D., & Crosas, M. (2020). *Advancing computational reproducibility in the Dataverse data repository platform* .

ukdataservice. Retrieved 3-11-2024 from https://ukdataservice.ac.uk/

University, H. *Information Security Policy Harvard University*. Retrieved 27-7-2024 from https://policy.security.harvard.edu/

University of Cambridge Research Data Management Policy Framework. (2021). Retrieved 27-7-2024 from https://www.data.cam.ac.uk/university-policy

University of Oxford Policy on the management of research data and records. University of Oxford .

Vejvoda, B ,.Burpee, K. J., & Lackie, P. (2017). Image Management as a Data Service. *IASSIST Quarterly*, *40*(3), 27. https://doi.org/10.29173/iq399

Vijai, D. C., Natarajan, D. K., & Elayaraja, M. M. (2019). Citation Tools and Reference Management Software for Academic Writing. *GIS Business*, *14*(6), 586-596. https://doi.org/10.26643/GIS.V14I6.14788

What is the most appropriate license for my research. figshare. Retrieved 27-7-2024 from

Zotero | About. https://www.zotero.org/about/

Prototype for establishing national research data repository

Mohamed Mosaad Mohamed Elzlabany

Library of Congress Cairo office

mh_zalabany@hotmail.com

Abstract

This research examines the legal frameworks and agreements governing the establishment of a research data repository at both international and national levels, alongside the ethical standards and institutional policies that regulate the management and sharing of research data. It provides a comprehensive analysis of data management policies, beginning with data collection, organization, and storage, and extending to data access, security, backup strategies, and other essential measures to ensure effective data protection and governance.

The study further explores the technological and infrastructural requirements necessary for establishing a national research data repository, drawing upon Harvard University's infrastructure in the implementation of the Dataverse system as a guiding model. It also reviews research data management systems and supporting software, as well as the operational requirements needed to sustain their efficiency and continuity. In addition, the research addresses repository configuration and control policies, the application of metadata standards, and the evaluation of repository performance through the development of indicators to measure both effectiveness and efficiency.

In conclusion, the research proposes a practical framework for the management of a national research data repository, encompassing financial, operational, and administrative planning to ensure the repository's long-term efficiency and sustainability. This framework is designed to achieve the overarching objectives of supporting scientific research and promoting the sharing of research data at both national and international levels.

Keywords: Research Data Repositories , Open Science Movement , National Research Data Plan