

الفصل الرابع

# ضبط الجودة الإحصائية لمستودع بيانات إلكترونية عربية\*

د. محمد مجدى أحمد قابيل

استاذ مساعد بقسم الحاسب الآلى ونظم المعلومات

أكاديمية السادات للعلوم الإدارية

---

\* أصل العمل مقدم باللغة الإنجليزية.



تعرض الورقة المقدمة تقريراً عن دراسة بحثية لتطوير منظومة ضبط الجودة الإحصائية لمستودع بيانات الكرونية عربية. حيث تعتمد كثير من القرارات المعاصرة على البيانات الموجودة في مستودع البيانات هذا. وتعتمد سلامة هذه القرارات إلى حد كبير على جودة البيانات المستخدمة. ويستخدم مستودع البيانات الإلكترونية العربية قواعد بيانات كبيرة باللغة العربية تحتوي على ملفات يتراوح حجمها ما بين ٥٠ ألف سجل وحتى ملايين السجلات. وحساب الحروف والأرقام يصل حجم بعض الملفات إلى ٦٠ مليون حرف وحتى ١٢ بليون حرف. وقد أدى حفظ هذا الكم الكبير من البيانات مقترناً بالمعدلات العالية للوصول والتحديث بها إلى ضرورة بناء منظومة محكمة لضبط الجودة الإحصائية للبيانات.

وقد كان النظام السابق لضبط جودة البيانات في مستودع البيانات الإلكترونية العربية يبنى على أسلوب الفحص الكامل للبيانات بمراجعة ١٠٠٪ منها. حيث تأتي البيانات إلى المستودع من مختلف الإدارات المعنية في شكل كشوف مطبوعة على ورق. ويقوم المستودع بإدخال البيانات من الكشوف المطبوعة إلى الملفات الإلكترونية للحاسب الآلي، ثم يقوم باستخراج نسخة مطبوعة من البيانات المسجلة فعلاً على ملفات الحاسب، وترسل النسخة المطبوعة إلى الإدارة المختصة لمراجعتها بمعرفتها والتوقيع على اعتمادها وإعادة ثانياً، حيث يقوم المستودع بإجراء التحديث لتصحيح البيانات بملفات الحاسب وإعادة الكرة حتى الوصول لاعتماد الإدارة المختصة لكشوف البيانات كما هي. وطبقاً لهذا الأسلوب تعتبر الإدارة المختصة بإصدار البيانات هي المسؤولة عن سلامة هذه البيانات. ولقد أدى هذا الأسلوب إلى نقل جزء من عبء المسؤولية من على كاهل مستودع البيانات الإلكترونية العربية، لكنه في الحقيقة لم يساعد كثيراً على رفع الجودة الفعلية للبيانات المستخدمة.

ولتحقيق مستوى جودة مستهدف للبيانات المستخدمة في مستودع البيانات الإلكترونية العربية، فلقد كان من اللازم تطوير دورة جديدة للضبط الإحصائي ودمجها في منظومة العمل السائدة. وتعرض الورقة لأحدى الدورات الجديدة للضبط الإحصائي والتي تم بناؤها في مستودع البيانات الإلكترونية العربية. وتشمل الدورة على توزيع العينات والفحص والحسابات الإحصائية ومخططات السيطرة. كما تعرض الورقة بعض نماذج من نتائج الفحص والحسابات الإحصائية الخاصة بها.