Recordkeeping for project management information system in public procurement: an action research

Records Management Journal

Recordkeeping for project management information system in public procurement: an action research

Massimo Rebuglio, Filippo Maria Ottaviani and Alberto De Marco
DIGEP, Politecnico di Torino, Turin, Italy

Received 1 October 2023
Revised 6 October 2024
20 November 2024
1 December 2024
8 January 2025
Accepted 23 January 2025

Abstract

Purpose – This study aims to examine the impact of laws, regulations and standards (i.e. the sources) on implementing a recordkeeping system in the public sector. It then provides an approach to set up a recordkeeping system in public procurement that complies with modern project management (PM) paradigms, information technology best practices and sources.

Design/methodology/approach – This study used the action research methodology to implement a recordkeeping system module within a PM information system for an Italian public procurement project. The study involved mapping the existing information flows, analyzing stakeholders' needs and constraints and developing, testing and analyzing PM approaches and tools.

Findings – Regarding the impact of the sources, national laws set usability parameters and specify the content to store but often overlook storage methods and formats. Customs dictate data exchange formats, significantly affecting storage efficiency and retrieval, and are challenging to alter. Contracts outline content requirements and may standardize formats among parties. Standards establish foundational principles, ensuring records remain authentic, retrievable and usable over time. The proposed approach to address these constraints involves implementing information systems across multiple layers.

Practical implications – This research proposes a blueprint for effective recordkeeping in public procurement. This can lead to cost savings, reduced paperwork and faster access to critical data. In terms of proposed policies, it is also discussed how contracts could address these issues at the root by imposing suitable formats for PM data exchange.

Originality/value – The novelty of this study is associated with the proposed approach for designing and implementing recordkeeping systems that comply with laws, regulations and standards aligned with modern PM paradigms.

Keywords Electronic document, Records management systems, Information system prototyping, PMIS, Project management, Project management information system, Public sector, Records management

Paper type Research paper

1. Introduction

Good records management is essential in the public sector (Shepherd, 2015). Records management systems in public administrations are designed based on transparency and

Records Management Journal
Emerald Publishing Limited
0956-5698
DOI 10.1108/RMJ-10-2023-0052

accountability, both internally and for citizens, to build public trust (Chorley, 2017). In this context, data to be managed can be divided into two main categories: structured data (e.g. databases or spreadsheet tables) and unstructured data (e.g. invoices or e-mails) (Fleckenstein and Fellows, 2018).

Managing both structured and unstructured data in an integrated manner is a significant challenge for the public sector. For various reasons, public policies emphasize storing information in the form it was received, which is often unstructured (Fleckenstein and Fellows, 2018). However, structured formats are necessary to make data accessible. For example, the issue with an invoice is simple: "I would like to store the "PDF" of the invoice, but at the same time, I want my accounting software to store the relevant data in a structured format."

This study has two objectives. First, it explores the reasons for using unstructured data storage in the public sector. Specifically, it seeks to identify the sources – whether they come from standards, regulations, contracts, or customs – that require the retention of records in unstructured formats. Second, it proposes a straightforward approach to this issue, tailored for contexts where it is impractical to make changes to an existing records management system designed to store only unstructured data. Both objectives are approached using an inductive research methodology. Both the analysis of the sources and the development of the proposed approach were carried out during an action research project. The project focused on managing project data within a large-scale public procurement construction project in Italy.

1.1 Records management in construction public procurement

A Project Management Information System (PMIS) is software that supports project managers throughout the project lifecycle (Project Management Institute, 2017). A PMIS provides several software tools, including scheduling tools, work authorization systems and configuration management tools. It also integrates with databases, including corporate knowledge bases. In addition, the system automates the generation of reports about the project status and performance indicators (Yehorchenkova *et al.*, 2021; Van Besouw and Bond-Barnard, 2021).

Recordkeeping is a key feature of PMIS. Since a PMIS manages project data throughout the entire project lifecycle, effective recordkeeping becomes a critical function of a PMIS. What PMIS standards refer to as a "database" can be generalized as a persistence layer – the software component responsible for ensuring long-term data storage.

In the context of public procurement, laws, regulations, standards and customs (referred to as "sources") mandate the use of unstructured recordkeeping systems, such as Electronic Document and Records Management Systems (EDRMS). As discussed in this paper, these sources require that information in public procurement projects must be exchanged and stored in the form of "administrative acts" and other unstructured documents. Document management systems are used for this purpose, ensuring compliance with legal and regulatory requirements. Furthermore, efficient data handling is required, which means that the persistence layer must be managed electronically. Consequently, sources prescribe the use of an EDRMS for recordkeeping in public procurement (Rebuglio *et al.*, 2024).

In contrast, modern Project Management (PM) paradigms emphasize that PMIS recordkeeping systems should use structured information (i.e. in a tabular format). For example, the automatic generation of reports, which can vary in terms of target and level of detail, makes it difficult – if not impossible – to automatically process, filter and aggregate information when relying on unstructured data. A key feature of structured information is the ability to perform structured queries, which allow users to retrieve data from tables based on the content of specific cells.

A public procurement project management information system (P3MIS) is a specialized PMIS that assists project managers in public procurement. A P3MIS adheres to both the

prescriptions of the sources and the modern PM paradigms. In summary, a P3MIS is a PMIS that must operate under the constraints imposed by public procurement laws and regulations. Figure 1 illustrates how a P3MIS integrates into and aligns with the above disciplines.

This study addresses the challenge of designing and implementing a recordkeeping system for a P3MIS based on structured data while complying with laws, regulations, standards and customs. The research was conducted using an action research methodology, focusing on a public procurement construction project in Italy. The first objective of the study is to identify the sources that affect recordkeeping in a P3MIS. The second objective is to propose an approach that reconciles the prescriptions of the sources with modern PM paradigms:

RQ1. Which sources influence records management in public administrations, and how?

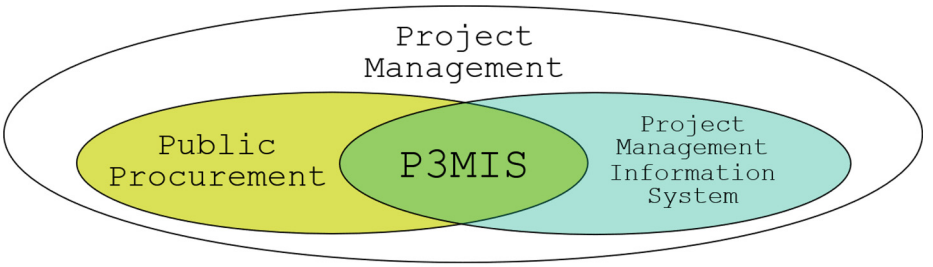
RQ2. How can a records management system be designed to enable structured queries on typically unstructured data?

This paper is organized as follows. First, we establish a common lexicon among laws, regulations, standards and the involved disciplines. Next, we review the relevant literature to identify the gap and to define the assumptions that justify the study – specifically the need to reconcile public sector data management requirements with those imposed by PM paradigms. Subsequently, we explain the concept of structured queries and why modern PM paradigms necessitate that a PMIS supports such queries. Then, we introduce the action research methodology and provide details about the research team involved in this study. Afterward, we present the project history, describing key interventions. Finally, we summarize how sources influence recordkeeping, underscoring the key points of our approach, and proposing generalizations.

1.2 Definitions

This paper uses the term *file* to refer to any document that contains information (e.g. a Word or a PDF file), while the term *structured data* to refer to a set of values associated with a specific field (e.g. name = Mario, surname = Rossi, birthdate = 01/ 05/72).

Let's consider the following terms used in the introduction: *document*, *record*, *(un) structured document and (un)structured information*. Since three disciplines (information technology, PM and recordkeeping) are involved, it is essential to establish a common terminology. According to ISO 15489–1 (2016) on records management, a *record* is defined as: "Information created, received, and maintained as evidence and as an asset by an organization or person, in pursuit of legal obligations or in the transaction of business." This definition implies that any information – regardless of its content or format – can be



Source: Figure by authors

Figure 1. Diagram illustrating how the P3MIS fits into the mentioned disciplines

classified as a *record* if it serves as evidence of business activities and is managed appropriately. The standard also specifies that *records* should have characteristics such as authenticity, reliability, integrity and usability and should rely on metadata to provide and preserve context and to apply appropriate management rules.

In contrast, in information technology, records are only the structured information (i.e. key-value pairs) while documents are generic information containers (The Editors of Encyclopaedia Britannica, 2024). In this context, a *record* is a single row in a tabular data structure, such as in a relational database, while *documents* are files. For example, an invoice is often a Word or PDF file, making it a document. Conversely, the structured data of an invoice, like customer name, date and total, can be stored in a tabular format. In this setup, each row of the table represents a *record* of an invoice. The terminology related to document-oriented databases (a kind of non-relational databases) is an exception: in these fields, a *document* is *structured information* organized as key-value pairs but without a tabular structure, meaning that not all records have the same keys.

Records management adopts yet another terminology, in which *documents* are mutable information and *records* are consolidated information (Adam, 2007). In this context, the difference between a *document* and a *record* is not in the format, but in the information status. *Documents* refer to information in the process of definition, subject to writing or approval procedures; *records* refer to consolidated immutable information.

In Table 1, the different terminologies are presented.

As mentioned at the beginning of this section, this paper adopts the following nomenclature to avoid confusion: *file* refers to *unstructured data*, *structured data* refers to data stored in *key-value* pairs. Also, this paper uses the term *structured query* to denote a search based on a key field, consistent with the definition of structured data (example: search for individuals born on 01/05/72).

2. Literature review

2.1 Integrity, authenticity and compliance in records management

The issue of record integrity and authenticity has been extensively explored in the literature, often in conjunction with efficiency and regulatory compliance. Duranti and MacNeil (1996) claim that while digital technologies have improved the production, transmission and storage

Table 1. Synoptic of the terminology in the disciplines and sources involved

Discipline/source	Document	Record
Standard (ISO 15489–1, 2016 on records management)		Information created, received and maintained as evidence of business activities; possesses authenticity, reliability, integrity and useability; relies on metadata
Information technology (the Editors of Encyclopaedia Britannica)	Generic information container (file), such as word or PDF files	Structured data (e.g. key-value pairs, a row in a table)
Document-oriented databases	Structured information as key-value pairs without a fixed schema	
Records management (Adam, 2007)	Mutable information in the process of creation or approval	Consolidated, immutable information serving as evidence

Source: Table by authors

of records, they have also increased the risks to the integrity and authenticity of those records, emphasizing the need to define requirements for managing and preserving them. A key aspect of Duranti and MacNeil's approach is the application of *diplomatic principles*. Diplomatics provides methods for analyzing the formal elements of records (e.g. structure, content and context) to determine their reliability. By applying diplomatic principles to electronic records, they argue that it is possible to preserve their value over time. These principles focus on ensuring that a record accurately reflects the event or transaction it documents and remains unaltered. In electronic systems, this means maintaining control over how records are created, stored and reproduced, especially since technological changes require ongoing processes, such as migration and reproduction to prevent obsolescence.

Cox (1997) builds on these concepts but shifts the focus beyond diplomatic principles, arguing that recordkeeping systems must go beyond simple storage and must comply with external regulations and authorities. According to Cox, an effective system should not only capture and maintain records but also ensure their long-term usability, which is critical for preserving their integrity and authenticity. Cox emphasizes that the reliability of records management systems largely depends on their ability to support continuous traceability and comply with regulatory requirements, thus, ensuring accountability. However, he also acknowledges the challenges many electronic systems have faced in maintaining record integrity, which challenges traditional definitions of what constitutes a record.

In the public sector, records management is particularly closely linked to the principles of transparency and accountability. Farrell *et al.* (2023) emphasize that access to accurate and reliable information is essential to ensuring the transparency of public authorities' actions. Without such access, citizens cannot effectively assess whether public institutions are operating in an accountable and transparent manner. However, despite the recognized importance of these principles, a significant gap remains between theory and practice.

Farrell *et al.* (2023) highlight that, while theories about access to information and the role of recordkeeping in promoting transparency are well-established, they do not always translate into practical, easily implementable solutions in public records management. Over recent decades, a global trend has emerged toward the adoption of Freedom of Information (FOI) or access to information legislation, driven by efforts to enhance transparency and accountability in governance (Ackerman and Sandoval-Ballesteros, 2006). In this context, the UK Government addressed these challenges by introducing the Freedom of Information Act (FoIA) in 2000. FoIA mandates public authorities to make records available upon request, thereby promoting transparency and public access to information. The Act also incentivized public bodies to adopt more efficient records management practices, including the use of electronic tools for document indexing and retrieval (Maguire, 2005).

This shift toward digitized records was intended to enhance the accessibility and management of public records, ensuring a more efficient flow of data and better storage and cataloging systems. Similarly, the European Union has made strides in promoting transparency and improving public records management. The EU's Regulation (EC) No 1049/2001 regarding public access to European Parliament, Council and Commission documents reflects a commitment to transparency across European institutions. This regulation enforces the principle that EU citizens should have the right to access institutional documents, ensuring accountability.

The Australian Government began working on DIRKS, a comprehensive manual for implementing a recordkeeping system (Lappin, 2010), mainly referring to file-oriented information, with a focus on the data flow and specifications for storage and cataloging. Recently, practices and tools for records management have been implemented and studied in emerging countries like Yemen (Mukred *et al.*, 2022), Botswana (Shonhe and Grand, 2020),

Namibia, Zimbabwe (Nengomasha and Chikomba, 2018; Shaanika *et al.*, 2018) and China (Yang *et al.*, 2020).

This work, however, refers to the principles of transparency not in terms of "open access" (i.e. making records available to citizens), but rather in ensuring transparency through a records management system that upholds the integrity of the processes themselves. For instance, it emphasizes the importance of storing data in the exact form in which it is received by public authorities, without any handling or alterations before it is recorded. We identified a few sources that address this specific issue, which we consider an innovative element of our work.

2.2 The issue of structured data in file-based contexts

Integrating document-based recordkeeping systems with structured data remains a significant challenge. As Duranti and MacNeil (1996) observed, incorporating metadata in file-based records management systems improves information accessibility. Therefore, the ability to automate tasks during or after archiving records is considered a sign of maturity for EDRMSs (Pappel *et al.*, 2022). Given that machine learning systems for natural language processing are inherently prone to errors, it is crucial to provide records management systems with structured data.

In several scenarios and sectors, there is a growing need to access aggregated data, which can only be achieved with structured data. Isa *et al.* (2015) examined multiple search functions within an EDRMS but did not include structured queries among them, highlighting a significant gap in current systems.

In some industries, this issue has been addressed by introducing standardized formats for the electronic exchange of specific information. For example, SAP uses IDocs – Intermediate Documents (SAP SE, 2024). IDocs are standardized data structures designed to enable data transfer between different systems. They facilitate the structured exchange of information, such as invoices, orders and delivery notes, allowing for the seamless flow of transactional data without manual input. This not only improves system integration but also reduces the likelihood of errors.

Despite these area-specific solutions, the literature review revealed the lack of standardized methods for managing structured queries within unstructured data. These limitations hinder the ability to perform complex data analysis and retrieval, which are becoming increasingly important in today's PM contexts (Warburton *et al.*, 2023).

2.3 Data management in public procurement project management information system

Data management in P3MIS is an important issue. PMIS are not just logical processors – they must manage project data efficiently so that stakeholders can view, update and refresh project reports in real time (Prifti and Dhoska, 2022). Several studies have shown that most project managers use inappropriate tools for data management, such as spreadsheets and paper-based documents (Archibald, 1992; Braglia and Frosolini, 2014). Even when well-designed PMISs are used to manage data consistently within an organization, significant challenges arise when integrating with external organizations (Van Besouw and Bond-Barnard, 2021; Braglia and Frosolini, 2014). The lack of standardized practices and guidelines often leads to frustration among staff, who improvise solutions to address arising problems (Caniëls and Bakens, 2012). In this context, public procurement represents an important case, as it often involves two parties – the public entity and the contractor – who must share project data.

PM relies heavily on structured data. Since the 1980s, researchers have documented methods for modeling project information using structured relational databases and

nonrelational structured databases (Ahlemann, 2009; Raymond, 1987) developed a generic PMIS model based on tuples (structured data) and tabular data. He further argued that designing a PMIS from a structured data model, rather than from its outputs, offers advantages in terms of scalability and agility. Ahlemann (2009) introduced a new data model for PMIS using unified modeling language (UML), a format for representing structured and relational data. Building upon the models of Schlagheck (2000), also based on UML, Froese (1998), which used the object-oriented paradigm for representing structured data, Ahlemann's work extended and refined these earlier approaches. Ottaviani *et al.* (2023) proposed a model grounded in structured and relational data that is suitable for data exchange formats and running simulations. In addition, Brziša (2011) proposed a model based on an XML schema, a format that is structured, albeit non tabular and leans more toward document orientation. In these cases, structured data allows for easy derivation of project metrics, fast and transparent information sharing and automation of the more routine parts of the PM process. However, none of these models address how to manage data storage in situations where a purely relational database is not feasible, and when flows of file-based information must be preserved.

3. Methodology

This paper adopts the action research methodology, a participatory and collaborative approach where researchers work alongside stakeholders to address real-world problems. The goal is to generate both immediate practical solutions and broader theoretical knowledge. Action research is characterized by cyclical and reflective planning, action, observation and reflection phases until the desired objectives are achieved (Kemmis *et al.*, 2014).

In this action research project, besides the authors – researchers from Politecnico di Torino – the Works Director of the contracting public administration, and the Technical Accounting Director of the private concessionaire operator were also involved. Table 2 details the composition of the research team.

According to Greenwood and Levin (2007), action research is "social research that simultaneously satisfies rigorous scientific requirements and promotes democratic social change." This approach is widely used and recognized as valid in the field of information systems (Baskerville and Wood-Harper, 1998; Wastell *et al.*, 2004), as it allows complex

Table 2. Research team composition

Kind	Organization	Details
Academic	Politecnico di Torino	Researcher, specializing in PMIS research for public procurement. He also has the skills to develop prototype information systems
		Researcher, specializing in project management and agile systems
Practitioners	Building and sustainability directorate of the research and teaching Centre (Direzione Edilizia e Sostenibilità)	Works director
	Contractor	Director of technical accounting

Source: Table by authors

problems to be addressed through direct interaction with the studied context. The fundamental principles of action research include: *active participation*, with the direct involvement of participants in the research process and operational decisions; an *iterative cycle*, consisting of a continuous process of planning, action, observation and reflection; and the *generation of practical and theoretical knowledge*, aimed to produce concrete solutions to problems and contribute to scientific knowledge.

Given the structure and social role of Italian universities, the authors believe that action research is particularly suited to support the digital transformation of public administrations. This methodology allows us to address the challenges of the Italian context, which is characterized by documentary and unstructured information flows, by promoting effective changes through direct interventions.

According to Kock and Lau (2001), action research practitioners, particularly in the context of information systems, must satisfy two masters: the first is the "client," in this case the contracting entity, which seeks immediate improvement to the current situation; the second is the scientific community, which requires methodological rigor and significant theoretical contributions. One might wonder why researchers chose action research over simpler and more established methodologies, such as surveys or observational studies. We provide two reasons: first, because of the social role of our universities in supporting public administrations; and second, because we believe certain insights cannot be obtained using traditional methods. For example, the case described in this paper involves operating in contexts where information flows are documentary, unstructured and hardly modifiable.

Anyone familiar with Italian public administrations knows this widespread scenario; however, it is not simple to trace its causes in a top-down manner, nor to identify how it could be changed. This situation arises from a mix of regulations, objective challenges, customs (which in Italy are a jurisprudential source) and inefficiencies. Setting up a survey to confirm a known fact would not be challenging, but asking individuals perfectly integrated into the system to identify its weaknesses would yield limited results. Action research, on the other hand, adopts a bottom-up approach, analyzing and deepening a single case that, although unique, represents a broader context; it does not merely analyze situations but actively intervenes with new ideas and skills, facilitating change and innovation.

3.1 Context and case selection

Within the public sector, public procurement is a significant context to examine records management practices. In terms of economic relevancy, public procurement represents 15–20% of global GDP (United Nations Economic Commission for Europe, 2019), and, especially in the construction sector, involves a substantial portion of the country's public expenditure. In Italy, it is managed by over 30,000 contracting authorities, including national ministries and regional bodies. Public procurement accounts for about 10% of Italy's GDP (European Union, 2016).

The case is representative of a large public procurement project in the construction sector. The case under consideration is a €160m investment in a public center for teaching and research, which is believed to be a representative example of any public procurement project where using PM techniques is worthwhile. As with each public procurement project, there is a public administration and a private contractor, and there are official ways and methods for exchanging information. The peculiarity of this project is that it is carried out using a fast-track approach, meaning, from the PM perspective, there is an overlap between the design and execution phases. From a data management perspective, this requires frequent and continuous updates to certain design documents. This last feature tends to strain the documentation system more than usual, but it does not alter its generalizability.

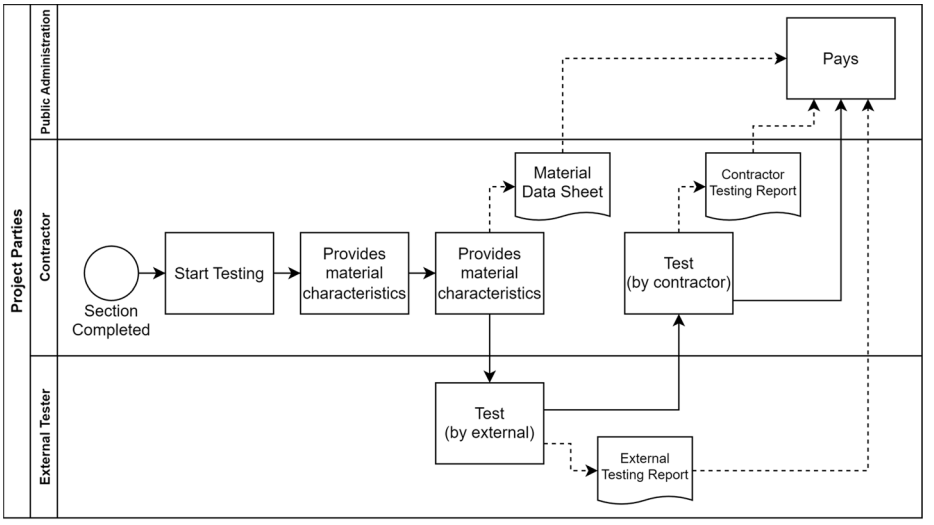
4. Project history

4.1 Context and scope

The intervention focuses on improving the process of monitoring the infrastructure project progress measurement. The contract defines the parts to be constructed: concrete pillars, pipes and tiles. The contracting authority must have evidence of this whenever a portion of the work is completed.

The as-is process is structured as follows: the tasks to be performed are divided into sections, each identified by a spatial code (indicating its location) and a technological-functional code (indicating what it pertains to). The private contractor, upon completion of work on a section, prepares one or more testing reports, possibly using external controllers and requests from its suppliers the specifications and certifications of the materials used. In addition, the operator marks the section as completed on a building information modeling (BIM) model, which tracks the progress in tabular and graphical format. However, despite the existence of the BIM model, the only legally binding evidence is the testing report, which must comply with a previously defined control plan for the specific technological-functional code. The Works Director needs to determine the progress status of either the entire project or a subset of it. Figure 2 provides the complete diagram of this process.

The EDRMS developed by the operator complies with standards and regulations, storing files and metadata in a key-value format. However, it does not allow for searches that involve the file content, and even less structured queries on the metadata. Another issue is the mismatch between the storage unit used to store the testing report on the EDRMS and the sections of work defined by the contract: a single testing report (or file) might pertain to more than one section, or a single section might require multiple testing reports. Due to these complexities, the checks on the progress status are substantially manual, with the Works Director's staff comparing the control plan with the sheets uploaded for each section. While this procedure was manageable for the sections completed initially, like the few but large



Source: Figure by authors

Figure 2. Outline of information flow

concrete pours, it would soon become unsustainable for the thousands of finishes and installations.

4.2 Intervention

Various solutions were considered to address this issue. The first option involved using the BIM, which consists of structured and queryable data, to assess the project progress instead of relying on the set of testing reports. However, this turned out to be unfeasible due to the regulatory context. The Italian public procurement law (Decreto Legislativo 31 marzo 2023, n. 36) mandates a formal testing phase within public contracts. The project contract further reinforces this requirement by stipulating that the control documentation is a prerequisite for making payments.

Customs also play a key role and deserve separate consideration. While customs are not as rigid as laws or contracts, they are among the most complex sources to deviate from. Specifically, the custom is to produce much of the required documentation, indeed, as documentation, meaning as unstructured files. It is unlikely that a tile supplier will provide its certifications in a structured format or agree to insert them into a shared structured database. They would simply send an e-mail with a PDF file containing the documentation.

Given the impossibility of solely relying on structured information, we thought about how to best index the existing unstructured data. It was proposed to introduce a new small computer tool, like a spreadsheet, that summarizes the unstructured files. This second approach, which at first glance seemed to make the system efficient without violating the sources, in practice, would have bypassed parts of it. The current EDRMS, in fact, strictly adopts the ISO standard on Access Control (ISO 15489–1, 2016), which would hardly have been possible on a shared table placed on parallel software. Also, reliability would have been lost, relying not on primary data but on a post hoc indexation, carried out in contradiction with the standard recommendations in terms of reliability: “Records should be created at the time of the event to which they relate, or soon afterward, by individuals who have direct knowledge of the facts.”

A summary of this information is described in Table 3, which provides a schematic overview of the sources and their impact.

Once the basic solutions were discarded, the one that later became the definitive solution was designed. It relies on a simple principle:

> Principle: The recordkeeping system must manage a structured datum for each file inserted. These two elements must always function as a pair in a one-to-one relationship. They are created, modified and deleted together.

By implementing a recordkeeping system that follows this principle, PM practices and source integrity would be maintained. However, the current system (and many EDRMS platforms, as noted in the literature review) was unable to handle this requirement. We then explored how to use the existing system effectively.

Our solution is based on two key assumptions:

(1) All information is contained and managed solely by the current file-based EDRMS, using existing authentication and reliability standards.

(2) Before being uploaded, files are enriched with structured data, which is embedded into them and becomes an integral part of each file.

This solution does not simply index existing content but enriches it. Following the two assumptions, there is a data retrieval approach inspired by non-relational databases: each file contained in the EDRMS is treated as a document (i.e. as the storage unit of a documentary database), and through a specific software, structured query interrogation can be performed.

Table 3. Summary of sources and impacts on recordkeeping

Source	Details	Impact
Regulations	Italian public procurement Law: "Public contracts are subject to testing for works and conformity checks for services and supplies, to certify that the object of the contract in terms of performance, objectives, and technical, economic, and qualitative characteristics has been realized and executed in accordance with the provisions and contractual agreements." * (Art. 116, Decreto Legislativo 31 marzo 2023, n. 36)	The system must store information related to testing
Contracts	Contract specifications: "quality control will consist of taking samples, performing tests on materials, acquiring certifications and provenance paths of the materials used, performing functional checks, operation tests, and provisional tests. Quality control will be the subject of specific programming to be formalized through the control plan. The existence [of such documents] is a sine qua non condition to accounting." * (contract of the procurement)	The system must store information related to materials, functional tests and trials
Customs	Testing reports are carried out on site, on paper. Material data sheets are made by professionals on digital files, then sent by email or delivered by hand. Technical material sheets for materials are files	The system cannot simply be a structured database, but rather must contain files

Note: * Translation from Italian
Source: Table by authors

4.2.1 Implementation details. In the specific implementation, we used the structured data format JSON. JSON (JavaScript Object Notation) is a lightweight, easy-to-read and easy-to-write format for structuring data. It represents data as key-value pairs, making it simple to transfer between systems, especially between a server and a client. Each piece of information has a "key" (the label) and a "value" (the data). For example, information in Testing Reports can be represented by a string like the following one: {"Type": "SC – Control Sheet", "Discipline": "SE – External works", "TechnologicalFunctionalPBS": "A1012 – Foundations", "SpatialPBS": "CS.110.3 – Ground Floor"}, where the specific tested element is included, identified by a functional technological key (for example, concrete column, copper pipe) and a spatial one.

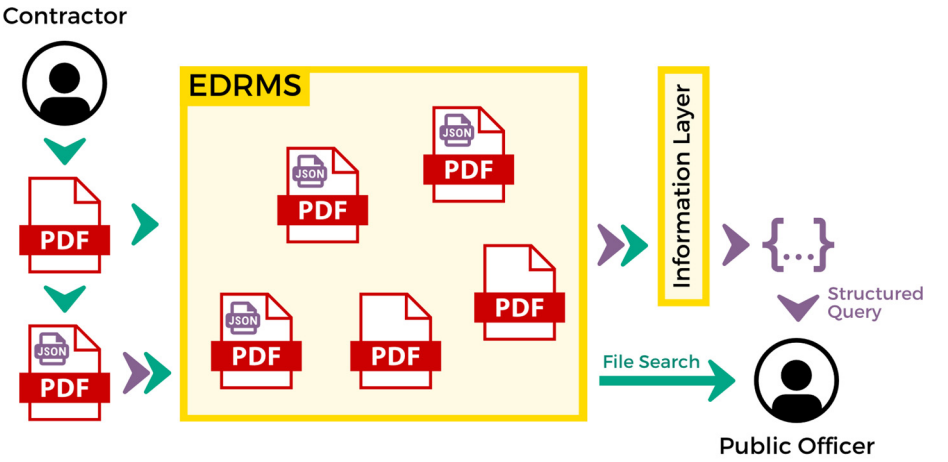
The previous principle was then implemented following two propositions:

P1. To make a file-based EDRMS queryable with structured queries, a correspondence between EDRMS and a document-oriented database can be defined, with a 1:1 correspondence between file and JSON document.

In fact, from a practical point of view, not wanting to change the existing software at all, two support applications were implemented, one responsible for inserting metadata on the documents, the other responsible for dispatching the structured queries. The first simply accepts PDF files and inserts a JSON string as an overlay and in the metadata. The second, more complex, ensures the possibility of performing structured queries:

P2. To avoid data replication, the Electronic Document Management System should be used as a persistence layer (instead of another structured database).

Figure 3 illustrates the proposed approach. It shows how each PDF is associated with structured data (specifically, a JSON string embedded within the metadata) before being uploaded to the



Notes: EDRMS ensures compliance with the sources; the information layer provides access to the structured data required by modern PM paradigms; green arrows represent an unstructured data stream, purple arrows a structured data stream
Source: Figure by authors

Figure 3. Outline of the proposed solution

EDRMS. From the EDRMS's perspective, the structured data becomes an integral part of the document, ensuring that any edits or deletions immediately apply to both the document and its associated structured data. An information layer has been implemented, using caching policies to extract the structured data and make it accessible through structured queries.

For testing reports, the structured data inserted consists of three key-value pairs: a technological-functional identifier (e.g. cement column, water pipe), a location identifier and a document type identifier (e.g. internal or external test report). This allows the extraction of information with a single query, such as the percentage of test reports completed for a specific area or for a particular technological-functional type. This system helps track progress and also determines when a section has been fully tested, signaling that payment can proceed, as indicated in Figure 2.

4.3 Evaluation

The proposed solution separates functional logic from the persistence layer and aligns with Braglia and Frosolini (2014), modeling a shared "technical database" among the parties involved in the project.

We believe our model easily integrates with the implementations of each theoretical model described in the literature review, albeit with specific precautions. For instance, many of these models frequently require data to be both structured and relational, something a document-based database does not provide. In such situations, one can apply the "external id" technique, which means referencing a unique identifier from another document within a given document.

The solution was generally well received by the parties involved:

- The public administration now has access to structured data for project monitoring and third-party checks, while still adhering to the document management requirements mandated by the sources.

- The private entity also gains access to the structured data for monitoring purposes, which is useful for conducting first-party checks. Importantly, they did not change their previously established document management system. As also highlighted by Maguire (2005) regarding metadata insertion, the process of embedding structured data into the document proved to be a critical phase: the designated operators viewed this step as an added burden. This issue was resolved by demonstrating how this approach was, in fact, a simplification. It required the production of structured information only once, at the inception of uploading to the EDRMS, rather than countless times afterward.

In addition, approximately one and a half years after the collaboration ended:

- The solution is still in use. Both the public administration and the contractor are in contact with the research team to address technical issues. Nonetheless, the objective of the collaboration was not to deliver a commercially viable solution but to assess the feasibility of the proposed approach, which was implemented with very limited technical and human resources.
- The main technical issues were due to EDRMS not "cooperating" in providing files to the information layer. To address this, a "bot" was implemented as a workaround to simulate user interaction and download the files. However, this portion of the code proved to be fragile, resulting in several problems during the file reading.
- A positive relationship with the contractor was essential in implementing this solution. The contractor's staff entered structured data into the files using our application, despite this not being required by the contract. This unexpected yet essential contribution suggests that, when drafting procurement contracts, it is important not only to specify the data to be exchanged but also the formats, favoring natively structured formats or, as in this case, "hybrid" ones.
- In general, we believe that natively structured solutions (e.g. SAP's IDoc) are preferable. These solutions offer the advantage of providing users with familiar, document-like interfaces while storing data in a natively structured way. Two considerations arise here from this: first, the contracting phase is critical, as the format must be agreed upon if necessary. On the other hand, in the specific case of public procurement, we believe that legislators should adopt policies to promote the development and adoption of such formats. An equivalent of IDoc for public works monitoring would undoubtedly have a significant positive impact on the quality of monitoring and control.
- The last two points of this example seem to "contradict" the conclusions of the first part of the study: in effect, we propose replacing formats typically imposed by custom with those enforced through contract terms.

5. Implications

From a theoretical perspective, the study deepens our understanding of the profound influence that laws, regulations, standards and customs (referred to as the sources) exert on recordkeeping in public procurement. By emphasizing the importance of multi-layered system designs to meet the various requirements posed by these sources, the research broadens scholarly debates on system design, particularly in contexts heavily shaped by external directives. More broadly, the strategies and some of the outcomes we have presented can be applied in any context where there are domain-specific needs and constraints tied to

laws and regulations. Additionally, the study affirms the suitability of action research as a method for studying information systems.

In terms of practical implications, this research offers guidance for practitioners faced with the challenges of designing recordkeeping systems for public procurement. The study presents a blueprint for establishing systems that align with regulatory mandates while also incorporating modern PM paradigms. Through a comprehensive mapping of current information flows and a focus on usability parameters, content types and storage formats, this research presents insights crucial for effective recordkeeping in public procurement contexts. The developed and tested prototype further demonstrates the potential and reliability of the architecture proposed by the study, providing a reference point for organizations looking to implement P3MIS.

6. Conclusion

According to Baskerville and Wood-Harper (1998), action research can be considered successfully completed based on two elements. The first is the immediate resolution of the issue, and the second is the identification of a theoretical framework to contextualize the solution. We believe we have addressed the first criterion in the "Intervention" and "Evaluation" sections, where we detailed our proposed solution and its subsequent acceptance. Regarding the second criterion, we provide answers through our two main research questions.

Regarding *RQ1* (*Which sources influence records management in public administrations, and how?*), we have summarized the findings of our research below:

- *National Laws*. National laws are heterogeneous sources. They can drive efficient recordkeeping by imposing maximum terms for data processing and promoting digital transition (Maguire, 2005; Shonhe and Grand, 2020), i.e. *usability parameters* for the systems. However, as this research has shown, they also tend to dictate the *content* of the information to be stored without focusing on the methods and formats in which it is stored.
- *Customs*. Customs dictate the practical ways in which involved parties exchange information. They, therefore, define the *formats* in which data can be stored and, consequently, significantly impact the efficiency of storage and retrieval. Even though they are not formally written down, they are the most challenging to change, especially when the number of actors is high and when they belong to external organizations. Given that recordkeeping systems tend not to be universal, it is advisable to consider customs during their design. The fact that customs influence formats, that is, the media through which data is preferred (or required) to be handled, aligns with previous literature (Oliver and Foscarini, 2019).
- *Contracts*. Contracts, like national laws, can specify the *contents* to be stored. In the context of public procurement, they expand and detail the content of the national laws on the subject. They can also commit to using common *formats*, limited to the parties involved in the contract.
- *Standards*. Standards are a category unto themselves, distinct from other sources. They are "agnostic" concerning content and formats. In the realm of records management, universally accepted standards set foundational principles for the creation, capture and administration of records. These principles ensure that records remain authentic, retrievable and traceable over time, regardless of their form or medium. Standards emphasize the significance of records management as a crucial tool for administrative efficiency, fiscal responsibility, transparency, risk

management and the preservation of institutional memory. They set out requirements for authenticity, reliability and integrity, which collectively provide *guidelines for data treatment*. Moreover, they advocate for *usability requirements*, emphasizing that records should remain accessible and functional throughout their lifecycle.

In response to *RQ2* (*How can records management systems be designed to enable structured queries on typically unstructured data?*), we identified a key functional requirement. Specifically, the recordkeeping system needed to manage both the inserted file and a corresponding structured datum. These two elements had to be linked in such a way that they always moved together as a pair, maintaining a 1:1 relationship. This meant that they were created, modified and deleted simultaneously, ensuring consistency throughout the process. In the implementation adopted in the case study, this was done by inserting a JSON metadata within the PDFs before entry into the EDRMS and introducing a logic layer that used the EDRMS itself as the persistence layer.

6.1 Policy considerations

Our findings highlight several key policy considerations for enhancing records management in public administrations. A crucial takeaway is the necessity of specifying not only the data to be exchanged in procurement contracts but also the formats in which this data should be provided. Favoring natively structured formats or "hybrid" solutions can greatly improve data accessibility and usability. As seen in the Evaluation section, maintaining a positive relationship with the contractor was instrumental in implementing our solution. The contractor's willingness to provide structured data, even though it was not a contractual obligation, illustrates how cooperative engagement can yield significant benefits. However, it is unrealistic to assume that such favorable conditions will always be present. To ensure the effective exchange of structured information, it is necessary to intervene at the contractual level. In the case we analyzed, contracts specified the content but not the formats, which hindered the efficiency of data exchange. We believe that imposing format requirements in contracts, breaking away from traditional practices, is essential for efficient electronic data exchange.

Moreover, we advocate for the adoption of natively structured solutions, such as SAP's IDoc, which combines familiar user interfaces with structured data storage. Once again, the contracting phase is critical, as agreeing on standardized formats can significantly improve interoperability and data consistency.

In this regard, legislative intervention is also required in public procurement. Policymakers should consider promoting the development and adoption of standardized formats to improve the monitoring and control of public works and enhance overall records management practices.

6.2 Limitations and generalizability

The paper adopts an inductive methodology, and consequently, the evidence presented is primarily valid for the case described. We believe three levels of generalization are possible, in increasing scope:

(1) In other Italian PMIS in public procurement. Our results are valid in any other case of this type. In the "Case Selection" section, we explained why we consider this case study to be a good representation of the phenomenon.

(2) In other Italian public contexts. The assumptions can be extended to other sectors of public employment where there is a need to manage both documentary and

structured data. For instance, we believe that the sources mentioned in *RQ1* similarly affect systems that PMIS should normally integrate with, such as accounting software (Ahlemann, 2009) and project resource management software (Van Besouw and Bond-Barnard, 2021); and that, as a result, the validity assumptions of the approach proposed in response to *RQ2* are confirmed.

(3) In other contexts outside the Italian regulatory framework. The results of *RQ1* are not valid outside the Italian context. However, the results of *RQ2* are generalizable, provided that the necessary conditions are met, namely the intention to maintain documentary data as the primary "source of truth."

References

Ackerman, J.M. and Sandoval-Ballesteros, I.E. (2006), "The global explosion of freedom of information laws", *Administrative Law Review*, Vol. 58 No. 1, pp. 85-130.

Adam, A. (2007), *Implementing Electronic Document and Records Management Systems*, CRC Press, Boca Raton, FL, doi: 10.1201/9780849380600.

Ahlemann, F. (2009), "Towards a conceptual reference model for project management information systems", *International Journal of Project Management*, Vol. 27 No. 1, pp. 19-30, doi: 10.1016/J.IJPROMAN.2008.01.008.

Archibald, D. (1992), "Understand the science before you apply the tools", *Aircraft Engineering and Aerospace Technology*, Vol. 64 No. 4, pp. 6-8, doi: 10.1108/eb037225.

Baskerville, R. and Wood-Harper, A.T. (1998), "Diversity in information systems action research methods", *European Journal of Information Systems*, Vol. 7 No. 2, pp. 90-107, doi: 10.1057/palgrave.ejis.3000298.

Braglia, M. and Frosolini, M. (2014), "An integrated approach to implement project management information systems within the extended enterprise", *International Journal of Project Management*, Vol. 32 No. 1, pp. 18-29, doi: 10.1016/j.ijproman.2012.12.003.

Brziša, S. (2011), "XML-based specification of the project management domain and its application", *Frontiers in Artificial Intelligence and Applications*, IOS Press, Vol. 224, pp. 213-226, doi: 10.3233/978-1-60750-688-1-213.

Caniëls, M.C.J. and Bakens, R.J.J.M. (2012), "The effects of project management information systems on decision making in a multi project environment", *International Journal of Project Management*, Vol. 30 No. 2, pp. 162-175, doi: 10.1016/j.ijproman.2011.05.005.

Chorley, K.M. (2017), "The challenges presented to records management by open government data in the public sector in England", *Records Management Journal*, Vol. 27 No. 2, pp. 149-158, doi: 10.1108/RMJ-09-2016-0034.

Cox, R.J. (1997), "More than diplomatic: functional requirements for evidence in recordkeeping", *Records Management Journal*, Vol. 7 No. 1, pp. 31-57, doi: 10.1108/EB027102/FULL/PDF.

Duranti, L. and MacNeil, H. (1996), "The protection of the integrity of electronic records: an overview of the UBC-MAS research project", *Archivaria*, pp. 46-67.

European Union (2016), "Public procurement – a study on administrative capacity in the EU".

Farrell, M., Gordijn, B. and Kearns, A.J. (2023), "Accountability, human rights and social justice in public sector recordkeeping", *Archival Science*, Vol. 23 No. 2, pp. 161-185, doi: 10.1007/S10502-022-09403-5/FIGURES/1.

Fleckenstein, M. and Fellows, L. (2018), "Records management", *Modern Data Strategy*, Springer International Publishing, Cham, pp. 195-207, doi: 10.1007/978-3-319-68993-7_17.

Froese, T.M. (1998), "Integrated computer-aided project management through standard object-oriented models".

Greenwood, D. and Levin, M. (2007), *Introduction to Action Research*, SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks CA 91320 United States of America, doi: 10.4135/9781412984614.

Isa, A.M., Ismail, W.M.H.W., Nordin, N.M. and Saman, W.S.W.M. (2015), "Information seeking behavior among IT professionals in using electronic records management system", *Advanced Science Letters*, Vol. 21 No. 6, pp. 2105-2108, doi: 10.1166/ASL.2015.6226.

ISO 15489-1 (2016), "Information and documentation — records management — part 1: concepts and principles", available at: www.iso.org/standard/62542.html (accessed 8 August 2023).

Kemmis, S., McTaggart, R. and Nixon, R. (2014), *The Action Research Planner: Doing Critical Participatory Action Research, The Action Research Planner: Doing Critical Participatory Action Research*, Springer, Singapore, doi: 10.1007/978-981-4560-67-2/COVER.

Kock, N. and Lau, F. (2001), "Information systems action research: serving two demanding masters", *Information Technology and People*, Vol. 14 No. 1, doi: 10.1108/ITP.2001.16114AAA.001.

Lappin, J. (2010), "What will be the next records management orthodoxy?", *Records Management Journal*, Vol. 20 No. 3, pp. 252-264, doi: 10.1108/09565691011095283/FULL/PDF.

Maguire, R. (2005), "Lessons learned from implementing an electronic records management system", *Records Management Journal*, Vol. 15 No. 3, pp. 150-157, doi: 10.1108/09565690510632337.

Mukred, M., Yusof, Z.M., Al-Moallemi, W.A., Mokhtar, U.A. and Hawash, B. (2022), "Electronic records management systems and the competency of educational institutions: evidence from Yemen", *Information Development*, Vol. 38 No. 1, pp. 125-148, doi: 10.1177/0266666920980829.

Nengomasha, C.T. and Chikomba, A. (2018), "Status of EDRMS implementation in the public sector in Namibia and Zimbabwe", *Records Management Journal*, Vol. 28 No. 3, pp. 252-264, doi: 10.1108/RMJ-08-2017-0023.

Oliver, G. and Foscarini, F. (2019), *Recordkeeping Cultures*, Facet, doi: 10.29085/9781783304011.

Ottaviani, F., Rebuglio, M. and De Marco, A. (2023), "Project management information system data model development and explanation ", *Proceedings of the 13th International Conference on Simulation and Modeling Methodologies, Technologies and Applications, SCITEPRESS – Science and Technology Publications*, pp. 210-217, doi: 10.5220/0012052200003546.

Pappel, I., Gelashvili, T. and Pappel, I. (2022), "Maturity model for automatization of service provision and decision-making processes in municipalities", *Lecture Notes in Networks and Systems*, Springer Science and Business Media Deutschland GmbH, Vol. 216, pp. 399-409, doi: 10.1007/978-981-16-1781-2_37.

Prifti, V. and Dhoska, K. (2022), "Information systems In project management and their role In decision making", *International Journal on Technical and Physical Problems of Engineering, International Organization on "Technical and Physical Problems of Engineering*, Vol. 14 No. 4, pp. 189-194.

Raymond, L. (1987), "Information systems design: a data modeling approach", *Project Management Journal*, Vol. 18 No. 4, pp. 94-99.

Rebuglio, M., De Magistris, P.E., Carlin, A. and De Marco, A. (2024), "Implementing EDRMS in public procurement: a retrofit approach", *Procedia Computer Science*, Vol. 239, pp. 541-546.

SAP SE (2024), "IDoc", available at: https://support.sap.com/en/alm/sap-focused-run/expert-portal/integration-cloud-monitoring/Idoc.html (accessed 5 October 2024).

Schlagheck, B. (2000), *Objektorientierte Referenzmodelle Für Das Prozess-Und Projektcontrolling*, Deutscher Universitätsverlag, Wiesbaden, doi: 10.1007/978-3-663-07993-4.

Shaanika, I., Maletzky, M. and Iyamu, T. (2018), "A model for addressing parallelism of systems in the management of data: the Namibian case", *African Journal of Science, Technology, Innovation and Development*, Vol. 10 No. 6, pp. 691-700, doi: 10.1080/20421338.2018.1491676.

Shepherd, E. (2015), "Freedom of information, right to access information, open data: who is at the table?", *The Round Table*, Vol. 104 No. 6, pp. 715-726, doi: 10.1080/00358533.2015.1112101.

Shonhe, L. and Grand, B. (2020), "Implementation of electronic records management systems: lessons learned from Tlokweng land board-Botswana", *Records Management Journal*, Vol. 30 No. 1, pp. 43-62, doi: 10.1108/RMJ-03-2019-0013.

The Editors of Encyclopaedia Britannica (2024), "Data structure", available at: www.britannica.com/technology/data-structure (accessed 8 January 2025).

United Nations Economic Commission for Europe (2019), "Sustainable procurement - minimal common sustainability criteria for sustainable procurement processes to select micro", *Small and Medium Sized Enterprise Suppliers.*

Van Besouw, J. V. and Bond-Barnard, T. (2021), "Smart project management information systems (SPMIS) for engineering projects – project performance monitoring and reporting", *International Journal of Information Systems and Project Management*, Vol. 9 No. 1, doi: 10.12821/ijispm090104.

Warburton, R.D.H., Ottaviani, F.M. and De Marco, A. (2023), "Critical analysis of linear and nonlinear project duration forecasting methods", *Journal of Modern Project Management*, Vol. 11 No. 1, pp. 188-199, doi: 10.19255/JMPM03113.

Wastell, D., Kawalek, P., Langmead-Jones, P. and Ormerod, R. (2004), "Information systems and partnership in multi-agency networks: an action research project in crime reduction", *Information and Organization*, Vol. 14 No. 3, pp. 189-210, doi: 10.1016/J.INFOANDORG.2004.01.001.

Yang, Q., Du, Y. and Shi, L. (2020), "Exploring the mechanisms for records management's digital transformation: a case study from China", *Records Management Journal*, Vol. 31 No. 1, pp. 34-47, doi: 10.1108/RMJ-10-2019-0064.

Yehorchenkova, N., Yehorchenkov, O. and Sazonov, A. (2021), "Project management information systems: an experience of developing and implementation on a production enterprise", *Case Study. In ITPM*, pp. 173-183.

Corresponding author
Massimo Rebuglio can be contacted at: massimo.rebuglio@polito.it