

A Socio-Technical Analysis of AI-Driven Legacy System Modernization in the Public Sector

Kumaresan Durvas Jayaraman

Independent Researcher, Seattle, State of Washington, USA

ABSTRACT

Purpose: This paper examines the socio-technical limitations of using artificial intelligence (AI) to modernize legacy information systems in the public sector, where continuity of services, policy obligations, security, and public accountability shape technology decisions.

Design/methodology/approach: The paper combines a review of literature on public-sector AI adoption and legacy-system modernization with an anonymous survey of more than 50 public-sector technology professionals. Participants represented technical, analytical, management, application-support, and project roles and a range of professional experience. Survey topics included AI benefits and risks, organizational readiness, security and privacy, governance, workforce challenges, and human oversight.

Findings: Respondents recognized AI's potential to support modernization but raised concerns about citizen-data protection, compliance, trust in generated outputs, integration, workforce readiness, and accountability. The findings suggest that AI should augment, not replace, experienced professionals and established engineering controls. Survey responses also point to the importance of governance, training, secure environments, phased implementation, and human validation.

Originality: The paper brings technical, organizational, institutional, and ethical constraints together in a public-sector legacy-modernization context. It emphasizes that AI should augment, rather than replace, accountable human expertise and that successful adoption depends on governance, secure data practices, validation, and organizational readiness.

Research limitations/implications: The survey reflects the views of its participants and should not be treated as representative of all public-sector technology professionals. Environmental impacts were not directly measured in the structured survey. Further research could test these findings across agencies and modernization projects.

Keywords: artificial intelligence; legacy system modernization; public sector; socio-technical systems; AI governance; human oversight

Article classification: Research paper

I. INTRODUCTION

What legacy systems are

Legacy systems are long-running information systems that were originally developed to support the core business operations of an organization. In many cases, these systems were built from the ground up using the technologies, infrastructure, programming languages, databases, and design practices available at that time.

These systems may still be stable and continue to support important business operations. However, technology keeps changing, and older systems may not easily support new business needs, modern security requirements, cloud integration, mobile applications, real-time services, or large-scale data processing.

Why governments still use them

Government systems often evolve over decades as new subsystems, rules, integrations, and operational processes are added. New subsystems, business rules, integrations, and operational processes were added whenever new requirements were introduced.

Recreating embedded business rules is difficult, particularly when documentation is poor and experienced staff are unavailable. [12] In public administration, successful transformation also requires a comprehensive understanding of the systems being changed, the policies they serve, and the broader administrative setting.

Rise of AI-based modernization

Modernization predates current AI tools: organizations have replaced legacy systems because technologies age, support ends, vulnerabilities emerge, and business needs change. AI adds assistance for code understanding, generation, translation, summarization, testing, and repair, but outputs still require engineering validation.

AI has now created a new level of support for modernization activities. [10] These capabilities can support modernization work, although their use still requires evaluation in the specific legacy-system context.

AI should support—not independently control—modernization. [7] In public-sector decision-making, human actors remain responsible for the use and implications of AI systems, including appropriate vetting, monitoring, and understanding of the tools they use.

Why government modernization differs from private sector

Government modernization is influenced by more than technical capability. [11] Research on public-sector AI likewise identifies social, economic, ethical, political/legal/policy, organizational, data, and technological challenges.

Government modernization must manage cost, schedules, security, compliance, procurement, accountability, and resistance to change. Unlike private firms, agencies cannot optimize solely for revenue or speed; continuity, equity, transparency, and statutory obligations also matter.

In the private sector, organizations may be more willing to allocate a dedicated budget and team where modernization can improve customer experience, reduce operating costs, increase revenue, or provide a competitive advantage.

Below are examples of legacy systems in the United States that support important public services and require high availability.

Unemployment Systems

Legacy unemployment systems process benefit applications, verify eligibility, calculate payments, and provide financial assistance to people who have lost their jobs.

Department of Motor Vehicles Systems

DMV systems support driver licensing, vehicle registration, title management, identity verification, and other public-safety and administrative functions.

Tax Systems

Government tax systems process tax returns, collect revenue, issue refunds, maintain taxpayer records, and support important financial operations.

Healthcare Eligibility Systems

Healthcare eligibility systems determine whether individuals and families qualify for public healthcare programs and related benefits. These systems process information such as income, residency, household details, and program requirements.

Social Welfare Systems

Social welfare systems support programs such as food assistance, housing support, childcare benefits, cash assistance, and other public services for eligible individuals and families.

II. BACKGROUND AND RELATED WORK

Legacy Systems

COBOL systems

COBOL systems continue to support many important government operations even though the technology was introduced many decades ago. [16]

Because these applications support important public-sector operations, modernization requires a clear understanding of the current system, its dependencies, business processes, and policy context.

Another challenge is the availability of technical resources. [12] Government modernization therefore needs deliberate

knowledge transfer and preservation of business and system knowledge.

Mainframes

Mainframe systems continue to provide a stable platform for many government applications. Government functions such as tax processing, unemployment benefits, pension systems, healthcare claims, payroll, and financial operations may depend on these environments.

Over time, legacy systems can become deeply interconnected with other applications, data, and administrative processes.

Government organizations may also face challenges related to aging infrastructure, limited availability of mainframe specialists, old integration methods, and incomplete documentation.

Tightly coupled architecture

Legacy systems can be tightly interconnected with other components and systems, so changing one area may affect others.

Government systems often become tightly coupled after many years of changes and new integrations. This can make maintenance, testing, enhancement, and modernization more difficult.

Undocumented business logic

Undocumented business logic is one of the major problems in legacy applications.

This is especially important in government systems because the hidden logic may support taxation, benefits, licensing, financial processing, or eligibility decisions.

AI-Based Modernization

Code translation

AI-assisted code translation can help developers transform source code into other representations or languages.

This can work well when the old and new applications have a similar structure or when the AI tool can access the related modules and dependencies.

The challenge becomes larger when a legacy application is being converted to a completely different technology or architecture. A legacy system may have many external dependencies, batch jobs, databases, file transfers, and separate applications that are not available in one codebase.

Therefore, AI-assisted code translation should be validated rather than accepted automatically. [7]

Automated testing

AI tools can support developers by generating test cases. [10]

AI can also help create integration tests and user-interface automation scripts using tools such as Selenium.

However, passing a large number of automated tests does not always mean that the complete business process is correct.

AI documentation generation

Creating and maintaining documentation is a large amount of work for development teams, and inadequate legacy documentation is a recognized modernization challenge.

AI can provide summaries of modules, interfaces, application flows, APIs, and business processes.

Another advantage is that the same information can be summarized in different ways. However, all generated documents should still be reviewed because AI may misunderstand incomplete or unclear information.

Business rule extraction

AI may also assist teams in recovering business knowledge indirectly through code understanding and requirements analysis.

After that, additional information such as old documents, source code, database logic, and system configurations can be provided.

The AI tool should also be instructed to separate existing business rules from assumptions or suggested improvements.

Chatbot interfaces

Chatbot interfaces make AI easier to use because developers, business users, and other stakeholders can communicate with the tool using normal language.

One limitation of generative AI is that apparently plausible outputs can still be incorrect. [7] Because of this, chatbot responses should be reviewed before they are used for important decisions.

Security and privacy are also important in public-sector AI adoption. [20] Sensitive source code or personal information should therefore be handled under the organization's approved security and data-governance controls.

When developers use chatbot tools frequently, the same interface can also help business users provide more information about processes and business rules.

Predictive system analysis

Predictive system analysis uses historical data, system patterns, performance information, and analytical models to estimate future problems or operational needs.

Predictive analysis can also help with planning. It may support decisions related to infrastructure needs, product lifecycle, maintenance cost, modernization priority, and future budget requirements.

The accuracy of predictive analysis depends on the quality of the available data. Legacy systems may have incomplete historical information, inconsistent data, or missing documentation.

III. MAJOR LIMITATIONS AND RESTRICTIONS

A survey involving more than 50 public-sector technology professionals was conducted to understand the technical, organizational, institutional, security, ethical, and workforce challenges associated with AI-driven legacy-system modernization. The survey examined participants' experience with legacy systems and modernization projects, their exposure to AI tools, perceived AI benefits and risks, organizational readiness, cybersecurity and privacy concerns, governance requirements, workforce challenges, and expectations regarding human oversight.

The survey explored professionals' practical views on AI-modernization risks and restrictions. Respondents recognized potential benefits but identified privacy, security, reliability, integration, governance, workforce, and budget constraints.

Participant Selection

Survey participants included workplace colleagues, LinkedIn connections, and professionals from government agencies and private companies, including individuals in senior technology and leadership roles. To protect participants' privacy, all responses were collected anonymously.

Data Privacy and Security Risks

Public-sector and government applications directly handle large amounts of sensitive information, including Personally Identifiable Information (PII), financial records, benefit information, healthcare-related data, employee records, and other confidential information.

In public-sector AI and legacy-system modernization, PII cannot be provided to AI tools without proper security, privacy, authorization, and data-handling controls. Agencies may need to establish approved AI environments, data-classification rules, access restrictions, security reviews, legal approvals, monitoring processes, and recovery procedures.

The risk becomes more serious when real citizen information is involved. Unlike test data, real personal information represents actual individuals and cannot simply be replaced if it becomes publicly exposed.

Data privacy and cybersecurity were major concerns identified in the survey. Approximately 92% of respondents identified citizen data protection as a significant concern, while 69% agreed that cybersecurity requirements make AI adoption more complex.

These findings support strict controls over data access, approved environments, data minimization, anonymization where feasible, and security review before AI tools are used with sensitive information.

Environmental and Resource Consumption Concerns

Another concern related to AI adoption is the amount of computing resources required to operate large AI systems. [13] These resource demands create legitimate sustainability questions around energy use, water consumption, carbon impact, and responsible AI deployment.

The survey did not directly include specific questions about environmental concerns such as electricity consumption, water usage, carbon emissions, or GPU resource requirements.

This does not mean that environmental concerns were considered unimportant. Approximately 62% of respondents considered cloud infrastructure important for AI-based modernization, showing that infrastructure requirements are relevant even though their environmental impact was not directly measured.

Future surveys could examine this topic more directly by asking public-sector technical professionals whether energy usage, water consumption, carbon impact, cloud-resource requirements, and infrastructure cost influence their views about AI adoption.

Limited Trust, Transparency, and Explainability

Limited trust in AI-generated results was one of the strongest concerns identified in the survey.

Only approximately 8% of respondents indicated that technical teams trust AI-generated outputs, while 85% identified low trust in AI tools as an organizational challenge.

These findings show that technical professionals are not prepared to allow AI-generated outputs to be used directly in critical government systems without verification. [18]

AI should support, not replace, accountable human decisions. Public-administration research emphasizes human responsibility, transparency, vetting, and monitoring [3].

Trust is more defensible when AI-supported decisions can be understood, reviewed, and justified.

Legacy-System Compatibility and Integration Challenges

Compatibility between modern tools and existing legacy environments is another major technical limitation. [11] Practitioners also report architecture, documentation, and legacy knowledge as modernization concerns.

Approximately 46% of respondents agreed that legacy-system architecture makes integration with modern platforms difficult, while another 46% directly identified integration difficulties as a major legacy-system challenge.

Other technical barriers were also identified. Approximately 31% of respondents identified limited infrastructure, 31% identified insufficient system documentation, and 23% identified a lack of integration capabilities.

Introducing AI into a legacy environment is therefore not simply a matter of connecting a model to a codebase. [7]

Teams must understand the existing architecture, dependencies, business rules, and operational context before selecting modernization tools.

Legal, Ethical, and Regulatory Restrictions

Legal, ethical, and regulatory obligations restrict public-sector AI use because agencies must protect citizens, justify decisions, retain records, and demonstrate compliance.

Approximately 85% of respondents identified regulatory compliance as a major public-sector constraint, while 92% agreed that AI-driven modernization must comply with transparency and accountability requirements.

Modernization is also affected by external changes, including policy priorities, procurement decisions, budgets, and leadership transitions.

Government AI adoption therefore requires attention to technical performance as well as organizational, ethical, legal, data, and policy conditions [20].

Approximately 85% of respondents supported establishing formal AI governance frameworks before AI is used in critical modernization projects.

A formal governance framework should define AI-use responsibilities, review requirements, and controls [20].

Dimensions of AI-Driven Modernization Constraints

A. Technical Limitations

1. Poor Documentation

AI-based modernization depends heavily on the quality of the information available in the existing system. [12] In public administration, successful transformation also requires a comprehensive understanding of the systems and the policies they serve.

Business rules are hidden

Business rules are often created when a system is originally designed and are based on the laws, policies, and operational requirements that exist at that time.

Another challenge is that older software projects may not have followed the same level of development standards and governance that are commonly expected today.

When AI is used for modernization or system rewriting, hidden or poorly documented rules create a serious validation problem. [7]

Outdated code lacks metadata

Older source code may also lack useful metadata, comments, descriptions, naming standards, and clear technical structure.

In some cases, improving the existing structure may be more difficult than rewriting part of the application.

For AI tools, the lack of meaningful metadata makes the problem more difficult.

Undocumented workflows exist

Poor documentation becomes even more serious when it is combined with complex workflows. During the original development of a system, there may have been situations where the technical team had to implement a solution quickly because of time, budget, resource, or project-delivery constraints.

In these situations, developers may have created complex workflows based on their understanding of the business requirement.

The modernization team must identify dependencies, external interactions, data movements, and exception paths before major changes are made. [7]

2. Data Quality Problems

Government applications collect and maintain large amounts of information over many years. While this information is important for public services and operational decisions, the quality of the data may vary depending on how the original application was designed and what validation rules were available at that time.

Poor-quality or incompatible data can complicate modernization. [11] Public-sector AI research also identifies data challenges as one of the major dimensions affecting AI adoption.

Incomplete records

Government systems may contain incomplete records when the application does not enforce enough validation during data entry.

For example, if an application collects a user's address but does not properly validate the address type, format, country, or

required fields, the system may contain information that cannot be reliably used later.

These issues may remain unnoticed for many years because the original application may continue processing the records.

Inconsistent formats

Inconsistent formatting is another common data-quality problem in legacy systems. As a result, users may have entered the same type of information in many different formats.

Examples can include phone numbers, names, date of birth, gender, addresses, identification numbers, and other personal information.

These inconsistencies can create significant difficulties during data migration. The modernization team may first need to identify, clean, standardize, and validate the existing data before it can be safely moved into a new system.

Decades of accumulated errors

When inconsistent and incomplete data continues to enter a system over many years, the errors can accumulate.

These accumulated errors become especially difficult when the organization introduces a new data structure, changes a database, or moves information into a modern application.

As a result, data cleansing and migration can become one of the most difficult parts of a modernization project.

AI outputs become unreliable

AI analysis is also affected by the quality of the data and artifacts supplied to the model. [20]

AI can help identify unusual values, repeated patterns, missing fields, and inconsistent formats, but the final results still require human review.

For this reason, AI should not be expected to automatically solve all data-quality problems.

3. Legacy Complexity

Legacy applications can become highly complex as they evolve over time. [11] Practitioners also report technical, business, and organizational challenges during legacy modernization.

This makes modernization difficult because the team must understand more than source code alone; system relationships, interoperability, policy context, and the broader administrative environment also matter.

Tightly coupled systems

Many older systems were developed for a specific purpose and were designed as independent applications.

Because of this design, sharing data or reusable services between applications can become difficult.

For example, if one government application needs information from another department or agency, direct sharing may be difficult when the systems were developed independently and do not use common interfaces or data standards.

Tight interdependencies create additional modernization work because relationships among systems and data must be understood before components are separated or replaced.

Custom protocols

Legacy systems may also depend on custom protocols, specialized interfaces, unique validation methods, and application-specific calculations.

Over time, different systems may have created separate versions of the same feature or process.

During modernization, technical teams need to decide which custom processes should be retained, which should be standardized, and which should be removed.

In many situations, large custom processes must be separated into smaller pieces before they can be standardized or reused in a modern architecture.

Obsolete programming languages

Technology continuously changes, and programming languages, frameworks, libraries, and development tools also evolve.

However, the issue is not always that the programming language itself becomes completely unusable. For example, applications written in languages such as C# or Java may still use the same basic programming concepts today.

When organizations upgrade the framework or move the application to a modern environment, parts of the source code may need to be rewritten or corrected.

Hardware dependency

Some legacy applications also depend on specific hardware, operating systems, device drivers, printers, scanners, or communication equipment.

An operating-system upgrade, infrastructure change, security-software update, or hardware replacement may affect how the legacy application performs.

For example, financial or banking environments may use specialized printers or other devices that communicate with mainframe applications.

Therefore, modernization planning should include hardware dependencies along with application and software dependencies.

AI cannot fully infer architectural intent

AI can help identify code dependencies, describe system components, and suggest architectural improvements.

Some architectural decisions may have been made because of policy requirements, hardware restrictions, security controls, performance needs, or limitations that existed many years ago.

Because of this, AI-generated architectural recommendations should be reviewed by experienced architects, developers, infrastructure specialists, and business experts before major changes are made.

IV. HALLUCINATION RISK

Generative AI can provide useful support during modernization, but its output is not always correct.

This risk becomes more serious in government systems because even a small, logical mistake can affect public benefits, financial transactions, eligibility decisions, taxation, or other important services.

Generate incorrect code

AI tools can generate code based on existing source code, documentation, user instructions, or prompts.

For example, an AI-generated program may compile successfully but still make an incorrect decision when approving a claim, processing an application, determining eligibility, calculating taxes, or creating an insurance transaction.

Many government business rules were developed over several decades and may contain special conditions, hidden meanings, manual adjustments, and exceptions.

Therefore, important AI-generated code should be tested and reviewed before acceptance. [7] Security evaluations of Copilot-generated code also show that generated suggestions can contain vulnerabilities.

Misunderstand policy logic

Policy-driven applications create an even greater challenge for AI-based modernization.

Examples include Paid Family and Medical Leave (PFML), Unemployment Insurance (UI), Disability Insurance (DI), and complex financial calculations.

When AI attempts to rewrite or modernize policy-driven systems, understanding source code alone may not be enough. [7]

As a result, an AI-generated solution can be technically plausible while still requiring validation against the governing policy and business context.

Introduce hidden defects

One major advantage of AI-assisted development is its ability to automate or assist a broad range of software-engineering tasks.

At the same time, faster code generation can also introduce hidden defects. Problems may appear only under specific business scenarios, unusual data conditions, or complex workflow paths.

These hidden defects can be difficult to identify unless the modernization team performs detailed testing across the complete application.

This does not mean that AI should not be used. Instead, the modernization project should include strong test planning, validation checkpoints, rollback options, version control, and recovery plans so that problems can be identified and corrected safely.

These concerns may appear complex and time-consuming, but they cannot be avoided in critical public systems. Similar validation would also be required in a fully manual modernization effort.

Examples of systems where hallucination or incorrect AI-generated logic can create serious impact include:

- Tax calculation systems
- Benefits eligibility systems
- Criminal justice systems

B. Security and Privacy Restrictions

Government legacy systems may contain sensitive information and support critical operations. [17] Public-sector AI research

also identifies data and political/legal/policy challenges as important dimensions of AI adoption.

Citizen Records

Public-sector legacy systems often contain sensitive citizen information such as personal details, addresses, financial records, health information, tax information, retirement-benefit details, licensing records, and legal records.

When generative AI is used during modernization, there is a concern that some of this information may be exposed to external AI platforms, processing environments, system logs, or other services.

Today, the same citizen information may exist in several government systems, but access is usually controlled separately within each agency or application.

Agencies should establish governance before allowing AI tools to process sensitive information. Controls should cover approved tools, access, data handling, retention, vendor terms, logging, security review, and human approval.

Social Security Numbers

Social Security Numbers are among the most sensitive pieces of personal information stored in government systems.

SSNs are already collected through secure government channels, but AI-assisted workflows may expose them to additional systems or vendors if data flows and access controls are not carefully designed.

Unlike traditional applications, AI-assisted workflows may transmit or reproduce sensitive data across tools, logs, or third-party services. Agencies must map those flows and minimize exposure.

Because of this, government agencies need strict controls when AI tools are allowed to access SSNs or other highly sensitive identifiers.

Health Information

Government organizations that provide health, social-service, family-support, and public-benefit programs may contain highly sensitive health and family information.

This information is highly sensitive and cannot be freely shared with external systems or other organizations without proper authorization and legal controls.

When AI is introduced into this environment, the same governance requirements become more complex.

For this reason, AI-based modernization involving health information should use approved and controlled environments. Access controls, encryption, data masking, authorization, auditing, and human review are important before sensitive health information can be processed by AI systems.

Law Enforcement Data

Law-enforcement systems contain information that is generally not available to the public.

Unauthorized exposure of this information can create serious consequences. Therefore, the use of AI tools with law-enforcement data requires a much higher level of security control.

Government agencies should use only approved and controlled AI environments when processing this type of information.

The concern is not that AI cannot technically process this information.

Overall Security and Privacy Restriction

Public-sector use of AI with government data requires controlled data and governance practices rather than unrestricted use. [3]

Security, data, legal, and policy concerns can therefore slow or constrain public-sector AI adoption.

C. Institutional Restrictions

Institutional restrictions are another important challenge in public-sector AI adoption. Empirical research identifies economic, political/legal/policy, organizational/managerial, data, technological, ethical, and social challenges that must be addressed alongside technical capability.

These processes are necessary because government organizations are responsible for using public funds and protecting public interests.

1. Procurement Policies

Government procurement processes often move slowly because they involve formal rules for budgeting, competitive bidding, vendor evaluation, contract approval, security review, governance review, policy review, and regulatory compliance.

These procedures are designed to promote fairness, transparency, competition, and accountability.

This becomes a larger problem with AI because the technology changes very quickly. An AI product selected at the beginning of a procurement process may have significantly changed by the time the contract is approved.

Procurement rules can delay small-scale AI trials. Agencies should use compliant pilot mechanisms, approved vendors, clear evaluation criteria, and safeguards for data and public funds.

Depending on the size and type of initiative, additional approvals may also be required at higher organizational or legislative levels.

At the same time, changing procurement rules is not simple. Therefore, government organizations need a balanced approach that supports innovation without reducing transparency and accountability.

Vendor evaluation is another time-consuming part of procurement. When only a few vendors are available, the evaluation may still require detailed technical, financial, security, and policy review.

A vendor may provide a strong technical product but still fail to meet specific government security, accessibility, privacy, data-retention, or procurement requirements.

Government agencies must also be able to explain and justify important uses of algorithmic systems. [3]

Agencies may therefore need to preserve prompts, generated outputs, source-code changes, approval records, testing results, access logs, and human-review decisions.

Although these audit requirements may slow the implementation process, they are necessary for transparency, accountability, legal compliance, and public trust.

AI product cycles move faster than government procurement. Agencies need procurement strategies that preserve fairness and oversight while enabling controlled pilots and timely reassessment.

2. Resistance to Change

People are one of the most important resources in the public sector.

Employee acceptance is an important part of AI adoption in government. [2], [9]

Resistance can be stronger when a new tool changes daily work habits. Many government employees remain within the public sector for a long period of time, and established processes can become part of their normal routine.

A similar situation happened when Agile development methods were introduced into information technology organizations. Many teams were originally comfortable with Waterfall-based development and were hesitant to move toward Agile methods.

Some employees may be concerned that AI will replace or substantially change their jobs. [2] At the same time, public-sector accountability research places responsibility for AI-supported decisions on human actors rather than on the algorithm itself.

AI should therefore be treated as an augmentation tool in high-stakes public-sector work rather than as a substitute for accountable human ownership.

Resistance to AI does not always mean that employees completely oppose the technology. As experience, training, governance, and confidence increase, the level of adoption may gradually improve.

Leadership support helps reduce resistance. Training, transparent communication, staff participation, and clear explanations of AI's role can build confidence [2].

3. Budget Constraints

Budget availability is one of the main factors that influences government modernization initiatives.

In many government organizations, budgets are prepared for a fixed financial period or biennium.

AI-driven legacy-system modernization can require investment in cloud infrastructure, AI platforms, software licenses, security controls, consulting support, employee training, and highly skilled technical professionals.

Hiring AI specialists can also be challenging in government organizations. Private-sector salaries for highly experienced AI professionals may be higher than the salary ranges available within established government job classifications.

This can make it difficult for agencies to compete for specialized AI talent.

Training existing employees also requires funding. Agencies may need to pay for training programs, development environments, cloud resources, software tools, and time for employees to practice and develop new skills.

At the same time, agencies must continue funding their existing systems and daily operations.

As a result, limited budgets can lead to delayed modernization, reduced project scope, insufficient employee training, or limited investment in security and governance.

A phased modernization approach may help manage these financial limitations. Clear cost-benefit analysis and long-term funding plans can also help leadership make better investment decisions.

4. Regulatory Constraints

Public-sector use of AI carries strong accountability expectations. Accountability requires actors to be able to explain and justify their conduct, and public-sector AI systems create additional challenges in maintaining that accountability.

When AI is used to modernize legacy systems, these requirements create additional work.

Black-box AI models create particular concerns in public-sector decision-making. Accountability research recommends preference for transparent and interpretable tools that allow public administrators to understand and justify their reliance on AI outputs.

If agencies do not preserve records of AI inputs, outputs, approvals, testing activities, and human reviews, they may not be able to demonstrate accountability or compliance later.

Public-sector AI processes should be traceable.

Human review should remain an important part of high-stakes public-sector AI use.

Regulatory requirements may slow the modernization process, but they are necessary because government systems directly affect citizens.

D. Ethical and Legal Limitations

AI-driven modernization can introduce ethical and legal concerns when systems affect citizens' access to public services or high-stakes decisions. [3]

One major concern is the possibility of unfair or biased outcomes. Public-sector AI research identifies ethical and data challenges as distinct dimensions of AI adoption, including concerns about how AI systems affect people and how data are used.

For example, an AI tool may correctly translate the technical structure of an eligibility calculation but still miss an important exception or policy condition.

Accountability is another major concern. Public-sector AI can create gaps in who can explain, justify, and be held responsible for algorithmic decisions, especially when multiple actors contribute to design, procurement, deployment, and use.

Even when multiple parties are involved, the use of AI does not remove human responsibility.

For this reason, AI-generated errors in critical government systems should not be treated only as software bugs or normal technical defects.

These risks can be reduced through controls that combine technical validation with organizational accountability. [7] Security evaluations show that generated code can contain vulnerabilities and therefore needs security review.

Explainability is important when AI contributes to significant public-sector decisions. Accountability research recommends transparent and interpretable tools so that public administrators can understand model functioning, potential failings, and the basis for relying on AI outputs.

Human approval therefore remains essential for high-stakes uses. Public-sector accountability research places responsibility on human actors, while software-engineering research recommends hybrid techniques that combine LLMs with conventional engineering methods.

AI should neither be blamed automatically nor excluded from modernization. Like code, data, infrastructure, and human decisions, AI outputs need governance, traceability, review, and clear responsibility for significant outcomes.

E. Workforce and Knowledge Limitations

Workforce and knowledge limitations are another major challenge in legacy modernization. [12] Public-administration research also stresses that teams must understand the systems, policies, and broader administrative setting in which legacy systems operate.

A current example is Agency Financial Reporting System, which serves as a central source of accounting information for State agencies and supports important financial activities such as payments, receipts, reimbursements, and connections with other budget and accounting systems.

Legacy financial systems have operated for many years and may contain large amounts of historical business logic, integrations, data structures, batch processes, interfaces, and agency-specific requirements.

The workforce challenge becomes more difficult when experienced legacy-system professionals are no longer available.

At the same time, professionals who specialize in AI, cloud computing, modern application development, cybersecurity, or modern integration technologies may not have deep knowledge of mainframe environments or government accounting processes.

This creates a rare combination of skills that government modernization projects need.

Agencies need multidisciplinary teams rather than a single person who knows every domain. Instead, agencies need to create teams where experienced legacy professionals, business subject-matter experts, architects, developers, AI specialists, cybersecurity professionals, and modern technology engineers can work together and transfer knowledge between each other.

The knowledge transfer is critical during large public-sector modernization efforts. If important legacy knowledge is lost before the new system has been completely understood, validated, and implemented, agencies may face data-conversion errors, incorrect financial processing, missed business requirements, integration problems, or disruption to important government financial operations.

AI can help with code analysis, code understanding, summarization, translation, testing, and related software-engineering activities. [10] However, model output can still be incorrect or hallucinated and should be combined with traditional engineering validation.

Effective modernization therefore requires both legacy-system knowledge and modern AI/software-engineering capability. [12] Government AI research highlights the importance of employee familiarity, training, and education.

Without this combination, AI may accelerate the technical part of modernization while still missing important business knowledge.

Framework Layer	Major Restrictions and Challenges
Technical	Legacy-system complexity, tightly coupled architecture, outdated technologies, poor documentation, hidden business rules, and integration difficulties
Organizational	Employee resistance, fear of job displacement, limited AI awareness, inadequate training, skill shortages, and loss of institutional knowledge
Institutional	Slow procurement processes, fixed budgets, restrictive policies, lengthy approval procedures, and limited flexibility for experimentation
Legal and Regulatory	Compliance obligations, explainability requirements, auditability, records retention, accountability, and responsibility for AI-generated errors
Ethical	Algorithmic bias, unfair treatment of citizens, lack of transparency, reduced human oversight, and uncertainty regarding automated decision-making
Security and Privacy	Exposure of citizen records, law-enforcement information, financial and healthcare data, cybersecurity threats, unauthorized access, and third-party AI risks

Layered Analytical Framework for AI-Driven Legacy System Modernization

The proposed framework organizes the major limitations of AI-driven legacy system modernization into interconnected technical, organizational, institutional, legal/ethical, and security/accountability dimensions.

Framework Contribution

The framework treats modernization as a socio-technical and institutional challenge, integrating technical reliability with workforce readiness, governance, security, ethics, and accountability. It supports checkpoint-based decisions about AI use, validation, human approval, and alternative approaches when outputs fail.

N o.	Moderniza tion Area	Government Agency Challenges	Private- Sector Approach	Governmen t Adaptation
1	Objectives	Public service, compliance, accessibility, security, transparency, and continuity take priority over profit.	Revenue, cost, customer retention, and competitiveness drive change.	Build a public-value case around service speed, accuracy, access, security, and citizen satisfaction.
2	Decision speed	Leadership, procurement, finance, legal, security, and legislative approvals can delay decisions.	Executives or product owners often decide quickly.	Delegate routine delivery decisions; retain oversight for major financial, legal, privacy, and policy matters.
3	Funding	Annual/biennial budgets limit mid-project flexibility.	Funding can shift with performance and priorities.	Stage funding across discovery, pilot, delivery, implementation, and expansion.
4	Procurement	Processes must ensure fairness, competition, transparency, and responsible spending.	Vendor testing, negotiation, and supplier changes may be faster.	Use modular contracts, prequalified vendors, small task orders, pilots, and performance-based agreements.
5	Delivery	Large sequential plans may lock requirements in years before delivery.	Agile, iterative, product-based delivery is common.	Release incrementally, prioritize high-value functions, collect feedback, and improve.
6	Risk tolerance	Failure may trigger audits, legal issues,	Controlled experiments may be	Run secure, limited pilots with

		legislative scrutiny, and public criticism.	accepted for business value.	synthetic data, human oversight, success measures, and stop criteria.			long projects.	ity.	, measurable goals, and shared ownership.
7	Replacing legacy systems	Tax, benefits, licensing, safety, finance, and health systems cannot tolerate prolonged outages.	APIs, parallel processing, microservices, cloud migration, and gradual replacement are used.	Avoid big-bang replacement; modernize capabilities incrementally and run old/new systems in parallel.	12	Data governance	Data may be fragmented, incompatible, and subject to different rules.	Common data models, APIs, platforms, and master-data practices are used.	Assign data owners, define authoritative sources, standardize formats, and secure interoperability.
8	Citizen needs	Services must include people with disabilities, limited internet, language barriers, or low digital literacy.	User research, analytics, journey mapping, and usability testing are common.	Use accessible, multilingual, assisted, telephone, and in-person options alongside digital services.	13	Vendor dependency	Proprietary contracts can create lock-in and limit access to knowledge, documentation, and code.	Providers may be changed and portable technologies adopted.	Require open APIs, documentation, portability, knowledge transfer, transition support, and code access where appropriate.
9	Workforce	Agencies may lose COBOL/main frame experts while struggling to hire AI, cloud, data, and security skills.	Private firms hire specialists, contractors, and competitive talent.	Pair legacy experts with modern engineers; mentor staff, document knowledge, and build internal capability.	14	Success measures	Success may be judged by launch, contract compliance, schedule, or budget.	Adoption, efficiency, reliability, and business outcomes matter.	Track processing time, availability, errors, effort, accessibility, incidents, satisfaction, and cost per transaction.
10	Organization	Technology, policy, legal, procurement, security, and program teams may operate in silos.	Cross-functional product teams are common.	Create teams spanning business, development, architecture, security, design, testing, and data.	15	AI and automation	Government AI must address privacy, bias, explainability, security, due process, records, and accountability.	AI supports development, service, forecasting, testing, documentation, and automation.	Start with code explanation, documentation, test generation, dependency analysis, and security scanning; retain human review for citizen-impacting decisions.
11	Leadership continuity	Political, executive, funding, and legislative priorities can change during	Modernization is tied to business goals and executive accountability	Maintain a long-term roadmap, documented governance					

Government vs. Private-Sector Modernization Approaches

This table compares major modernization challenges commonly faced by government agencies with the approaches generally used in the private sector. It also identifies practical ways that public-sector organizations can adopt selected private-sector practices without compromising legal, regulatory, security, accessibility, transparency, or public-accountability requirements.

Research Gap

Existing research covers several relevant streams separately: public-administration legacy transformation and interoperability, [11] LLM-based software engineering, [10] cybersecurity risks of federal legacy systems, [17] government-employee acceptance and training, [2] and the ethical, legal/policy, organizational, data, and accountability challenges of public-sector AI.

One important research gap is the lack of a clear framework that identifies the major checkpoints required throughout an AI-driven modernization project.

AI-generated code may accelerate technical work, but reliability still requires explicit validation. [3] These findings support the need for modernization checkpoints that test accuracy, security, policy alignment, and accountability.

Another area that needs attention is the social and organizational response to AI-driven modernization. [2] Broader public-sector AI research also identifies social, ethical, organizational, policy/legal, data, and technological challenges.

Research is therefore needed on how government organizations can prepare for these concerns before the modernization process begins.

Risk assessment is also an important research area. Public-administration legacy research shows that complexity and interoperability can challenge transformation, [11] federal evidence links legacy-system prevalence with security incidents, [17] and public-sector AI research identifies social, economic, ethical, legal/policy, organizational, data, and technology challenges.

The research gap is how government agencies can use AI safely and effectively while managing technical, institutional, human, legal, ethical, and security risks across the modernization lifecycle.

Areas Requiring Further Study

The following areas require additional research:

How can clear checkpoints be identified for each phase of AI-driven legacy-system modernization?

How do legacy complexity, poor documentation, and hidden business rules affect the reliability of AI-assisted modernization?

How should modernization teams respond when AI-generated code or recommendations fail validation?

How do workforce resistance, skill gaps, and loss of institutional knowledge influence AI adoption?

How do procurement processes, budget limitations, legal requirements, ethical concerns, and security policies restrict implementation?

How should government agencies communicate AI-related risks and changes to employees, leadership, business owners, and the public?

What project risks can AI-driven modernization efforts to stop, slow down, or change direction?

What governance, human-oversight, testing, and approval practices can support responsible AI-driven modernization?

How can government agencies combine legacy-system expertise, business knowledge, and modern AI skills within the same modernization team?

What type of reusable framework can help government agencies decide when AI should be used, when human intervention is required, and when an alternative modernization approach should be selected?

V. CONCLUSION

AI can assist with code understanding, translation, summarization, testing, and repair, but generated outputs may be incorrect and require validation [7], [10]. Government modernization must also account for legacy complexity, policy context, data quality, security, and continuity of public services.

This paper shows that the main challenge is not whether AI can support modernization, but how it can be used safely, responsibly, and effectively in a government environment. [17]

Human involvement remains essential. [3] Software-engineering research likewise recommends hybrid approaches to identify incorrect LLM outputs, [7] while security evaluations demonstrate that generated code can contain vulnerabilities.

Government agencies can also learn from private-sector modernization practices, including phased delivery, cross-functional teams, reusable services, API-based integration, controlled experimentation, and measurable outcomes.

Future research should develop a checkpoint-based framework that specifies when AI is appropriate, what evidence is required at each stage, when human approval is mandatory, and when teams should choose an alternative approach.

REFERENCES

- [1] Abu Bakar, H., Razali, R., and Jambari, D. I. (2022), "A qualitative study of legacy systems modernisation for citizen-centric digital government", *Sustainability*, 14(17), 10951. <https://doi.org/10.3390/su141710951>
- [2] Ahn, M. J., and Chen, Y.-C. (2022), "Digital transformation toward AI-augmented public administration: The perception of government employees and the willingness to use AI in government", *Government Information Quarterly*, 39(2), 101664. <https://doi.org/10.1016/j.giq.2021.101664>
- [3] Busuioc, M. (2021), "Accountable artificial intelligence: Holding algorithms to account", *Public Administration Review*, 81(5), 825–836. <https://doi.org/10.1111/puar.13293>

- [4] Chen, T., Gasco-Hernandez, M., and Gil-Garcia, J. R. (2026), "Understanding accountability in AI use within the public sector: Insights from state governments in the United States", *Public Administration*, 1–19. <https://doi.org/10.1111/padm.70051>
- [5] de Almeida, P. G. R., and dos Santos Júnior, C. D. (2025), "Artificial intelligence governance: Understanding how public organizations implement it", *Government Information Quarterly*, 42(1), 102003. <https://doi.org/10.1016/j.giq.2024.102003>
- [6] De Siano, G. D., Fasolino, A. R., Sperli, G., and Vignali, A. (2025), "Translating code with large language models and human-in-the-loop feedback", *Information and Software Technology*, 186, 107785. <https://doi.org/10.1016/j.infsof.2025.107785>
- [7] Fan, A., Gokkaya, B., Harman, M., Lyubarskiy, M., Sengupta, S., Yoo, S., and Zhang, J. M. (2023), "Large language models for software engineering: Survey and open problems", In *2023 IEEE/ACM International Conference on Software Engineering: Future of Software Engineering (ICSE-FoSE)* (pp. 31–53). IEEE. <https://doi.org/10.1109/ICSE-FoSE59343.2023.00008>
- [8] Fávero, L. F., de Almeida, N. R., and Affonso, F. J. (2025), "A systematic mapping study on the modernization of legacy systems to microservice architecture", *Applied System Innovation*, 8(4), 86. <https://doi.org/10.3390/asi8040086>
- [9] Haesevoets, T., Verschuere, B., and Roets, A. (2025), "AI adoption in public administration: Perspectives of public sector managers and public sector non-managerial employees", *Government Information Quarterly*, 42(2), 102029. <https://doi.org/10.1016/j.giq.2025.102029>
- [10] Hou, X., Zhao, Y., Liu, Y., Yang, Z., Wang, K., Li, L., Luo, X., Lo, D., Grundy, J., and Wang, H. (2024), "Large language models for software engineering: A systematic literature review", *ACM Transactions on Software Engineering and Methodology*, 33(8), Article 220. <https://doi.org/10.1145/3695988>
- [11] Irani, Z., Abril, R. M., Weerakkody, V., Omar, A., and Sivarajah, U. (2023), "The impact of legacy systems on digital transformation in European public administration: Lesson learned from a multi case analysis", *Government Information Quarterly*, 40(1), 101784. <https://doi.org/10.1016/j.giq.2022.101784>
- [12] Khadka, R., Batlajery, B. V., Saeidi, A. M., Jansen, S., and Hage, J. (2014), "How do professionals perceive legacy systems and software modernization? In *Proceedings of the 36th International Conference on Software Engineering* (pp", 36–47). Association for Computing Machinery. <https://doi.org/10.1145/2568225.2568318>
- [13] Li, P., Yang, J., Islam, M. A., and Ren, S. (2025), "Making AI less 'thirsty': Uncovering and addressing the secret water footprint of AI models", *Communications of the ACM*, 68(7), 54–61. <https://doi.org/10.1145/3724499>
- [14] Madan, R., and Ashok, M. (2023), "AI adoption and diffusion in public administration: A systematic literature review and future research agenda", *Government Information Quarterly*, 40(1), 101774. <https://doi.org/10.1016/j.giq.2022.101774>
- [15] Maragno, G., Tangi, L., Gastaldi, L., and Benedetti, M. (2023), "Exploring the factors, affordances and constraints outlining the implementation of artificial intelligence in public sector organizations", *International Journal of Information Management*, 73, 102686. <https://doi.org/10.1016/j.ijinfomgt.2023.102686>
- [16] Melo, I., Polónia, D., and Teixeira, L. (2025), "Human–AI collaboration in the modernization of COBOL-based legacy systems: The case of the Department of Government Efficiency (DOGE)", *Computers*, 14(7), 244. <https://doi.org/10.3390/computers14070244>
- [17] Pang, M.-S., and Tanriverdi, H. (2022), "Strategic roles of IT modernization and cloud migration in reducing cybersecurity risks of organizations: The case of U.S", federal government. *The Journal of Strategic Information Systems*, 31(1), 101707. <https://doi.org/10.1016/j.jsis.2022.101707>
- [18] Pearce, H., Ahmad, B., Tan, B., Dolan-Gavitt, B., and Karri, R. (2022), "Asleep at the keyboard? Assessing the security of GitHub Copilot's code contributions", In *2022 IEEE Symposium on Security and Privacy (SP)* (pp. 754–768). IEEE. <https://doi.org/10.1109/SP46214.2022.9833571>
- [19] Selten, F., and Klievink, B. (2024), "Organizing public sector AI adoption: Navigating between separation and integration", *Government Information Quarterly*, 41(1), 101885. <https://doi.org/10.1016/j.giq.2023.101885>
- [20] Sun, T. Q., and Medaglia, R. (2019), "Mapping the challenges of artificial intelligence in the public sector: Evidence from public healthcare", *Government Information Quarterly*, 36(2), 368–383. <https://doi.org/10.1016/j.giq.2018.09.008>
- [21] van Noordt, C., and Misuraca, G. (2022), "Artificial intelligence for the public sector: Results of landscaping the use of AI in government across the European Union", *Government Information Quarterly*, 39(3), 101714. <https://doi.org/10.1016/j.giq.2022.101714>
- [22] Zuiderwijk, A., Chen, Y.-C., and Salem, F. (2021), "Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda", *Government Information Quarterly*, 38(3), 101577. <https://doi.org/10.1016/j.giq.2021.101577>