



## Public Values in Saudi Arabian Artificial Intelligence Policy Documents: A Mixed-Method Analysis Using NVivo and the Bannister & Connolly (2014) Framework

Abdulrahman Rashed Dafer Althabiti

Department of Public Administration  
Faculty of Business Administration  
King Saud University  
Althabty99@gmail.com

### Abstract

*Background:* The integration of artificial intelligence (AI) into public sector governance generates complex value tensions that remain understudied in non-western institutional contexts. *Objective:* This study examines public values embedded in Saudi Arabian governmental AI policy documents, using the Bannister & Connolly (2014) framework as an analytical lens alongside qualitative content analysis conducted via NVivo 15. A purposive sample of nine official documents published between 2016 and 2024 was subjected to systematic matrix coding, generating 198 total coding references across three value categories. Two NVivo matrices were produced: a cross-tabulation matrix (Q3) yielding 222 pure diagonal references from 270 total, and a considerations-to-risks matrix (Q4) comprising 115 considerations and 54 risks ( $n = 169$ ).

Duty-Oriented Values dominate at 43.7%, followed by Socially-Oriented Values (28.4%) and Service-Oriented Values (27.9%), with a single-reference margin separating the latter two. The Q4 matrix reveals that Socially-Oriented Values carry the highest relative tension (ratio 1.33:1), while Duty-Oriented Values record the highest absolute risk count. Cross-contextual comparison with Toll et al. (2020) reveals a notable structural similarity: both Saudi and Swedish AI policy documents exhibit an identical 43.7% dominance of duty/professionalism values, which may suggest an institutional logic in which governments frame transformative technologies primarily through compliance and control lenses. Crucially, Saudi documents acknowledge risks at 32% of Q4 references versus 9.6% in the Swedish corpus, embodying the “more nuanced view” advocated by Toll et al.

The term second-generation governance is used here in the sense developed by Janowski (2015) and elaborated by Bullock (2019). Janowski (2015) maps four stages of digital government evolution, distinguishing first-generation models, characterised by the digitisation of existing services, from later-generation models that integrate adaptive policy-making, contextual responsiveness, and transformational governance objectives. Bullock (2019) extends this stage-logic into the AI policy domain by tracking the analytical shift from compliance-and-control framings of artificial intelligence (first generation) to dual-track framings that combine promotional optimism with explicit regulatory caution (second generation). Saudi AI policy documents, as the present analysis demonstrates, exhibit this dual-track structure: a substantial proportion of references articulate developmental ambitions while another substantial proportion of references articulate risk-acknowledging regulatory positions, the two coexisting within the same policy texts. Saudi AI policy documents represent a second-generation governance model that balances promotional optimism with legislative caution. The findings contribute to the emerging literature on AI governance values in non-Western institutional contexts and offer a replicable methodological framework for comparative document analysis.

**Keywords:** Artificial Intelligence Governance; Public Values; Saudi Arabia; AI Policy Documents; Qualitative Content Analysis; NVivo; Value Tensions.

\* This article was submitted in April 2026, and accepted for publication in May 2026. Published Online in May 2026.

DOI: 10.21608/aja.2026.486090.2086



## Introduction

The advent of artificial intelligence (AI) as a transformative force in public administration has generated an urgent scholarly question: which public values do governments embed in their AI governance frameworks, and how do those values hierarchies reflect or diverge from broader civilisational and institutional contexts? While a growing body of research has examined AI policy documents in Scandinavian (Toll et al., 2020), European, and North American settings, the Arab governmental context remains comparatively understudied in the public values literature.

This study addresses three interrelated research questions:

- (RQ1): What is the hierarchical ordering of public values in Saudi AI policy documents?
- (RQ2): What are the positive and negative implications of AI integration for each value category?
- (RQ3): How do Saudi governance documents compare with the Swedish context analysed by Toll et al. (2020) in terms of value distribution and risk acknowledgement?

By applying the Bannister & Connolly (2014) public values framework which categorises governance values into Duty-Oriented, Service-Oriented, and Socially-Oriented clusters alongside NVivo-assisted qualitative content analysis, the study produces both quantitative distributional data and rich qualitative interpretations of how Saudi governance articulates its relationship with AI.

## Literature Review

### *Public Values in Government and Digital Governance*

The concept of public values occupies a central position in public administration theory. Bozeman (2007) defines public values as normative consensus about the rights of citizens, the obligations of citizens to society, the state, and one another, and the principles on which governments and policies should be based. Bannister & Connolly (2014) extend this foundation to ICT-mediated governance, proposing a tripartite framework: Duty-Oriented Values (legal compliance, accountability, integrity, professionalism), Service-Oriented Values (efficiency, transparency, effectiveness, responsiveness), and Socially-Oriented Values (equity, fairness, privacy, public accountability). Crucially, this framework acknowledges that digital transformation does not merely implement values but actively constructs and contests them.

Central to this framework is the definition of value itself. Bannister and Connolly (2014, p. 119) define a value as “a mode of behaviour, either a way of doing things or an attribute of a way of doing things, which is held to be right,” where rightness is established through near-universal citizen consensus. This behavioural operationalisation distinguishes the framework from purely abstract philosophical treatments and enables systematic coding of governance documents.

A critical theoretical distinction underpins the framework. The concept of public value (singular), as articulated by Moore (1995), refers to the strategic output that public managers create through their organisations. Public values (plural), by contrast, denote the normative standards and principles that govern governmental conduct (Bozeman, 2007). Bannister and Connolly (2014, p. 120) explicitly connect these concepts, noting that “public sector values underpin both public value and good governance in a variety of ways.” The present study examines public values as normative standards embedded in policy documents, not public value as managerial output.

The tripartite classification maps onto Hood’s (1991) seminal typology of administrative values: Duty-Oriented Values correspond broadly to sigma values (economy and parsimony), Service-Oriented Values to lambda values (reliability and robustness), and Socially-Oriented Values to theta values (fairness and equity). However, Bannister and Connolly (2014, p. 123) deliberately extend each category beyond Hood’s original scope, incorporating 29 specific values across the three clusters: 10 duty-oriented, 6 service-oriented, and 12 socially-oriented values, plus whistleblower protection.

Furthermore, Bannister and Connolly (2014, p. 120) argue that “ICT impacts on the majority of public service values because it is both an enabler and an embedder.” This dual function is directly applicable to AI: as an enabler, AI creates new possibilities for value realisation (e.g., enhanced transparency through algorithmic explainability); as an embedder, AI encodes value choices into automated systems that subsequently constrain or shape administrative behaviour. The five dimensions of transformational impact identified by the framework -scale, scope, integration, speed, and capability (Bannister & Connolly, 2014: 124)- provide analytical categories for understanding how AI governance documents anticipate these impacts. In a complementary vein, Chen et al. (2023) observe that AI applications in the public sector generate value impacts across multiple governance dimensions, reinforcing the need for frameworks that capture the breadth of public values at stake.

Complementary frameworks by Rose et al. (2015) identify four governance value ideals professionalism, efficiency, service, and engagement that have been empirically validated across multiple European e-government contexts (Persson et al., 2017). The persistent dominance of professionalism/duty values across these studies suggests a structural logic in how public institutions approach technological transformation, one that the present study both confirms and extends to a non-Western context.

This deliberate cross-framework mapping rather than mechanical overlay is consistent with Pollitt and Bouckaert’s (2017) caution against forced framework imposition: each framework retains its distinct conceptual identity while their substantive overlap is identified through analytical reading rather than imposed equivalence.

### ***AI Governance and Public Values: Prior Empirical Evidence***

The landmark comparative study by Toll et al. (2020) analysed ten Swedish AI policy documents (n = 522 coded statements), finding that professionalism values dominated at 43.7%, followed by efficiency (30.1%), service (18.8%), and engagement (7.5%). Critically, the study identified an “overly optimistic” rhetorical disposition benefits outweighing risks at a 5.62:1 ratio and called for “a more nuanced view” of AI governance discourse. This finding provides the primary comparative benchmark for the present study.

Chadwick & May (2003) and Jæger & Löfgren (2010) observed that engagement and participatory values tend to receive limited attention in governmental technology discourse, despite stated commitments to stakeholder involvement. The present study notes a comparable observation in the Saudi context, which may point toward institutional dynamics that transcend cultural specificity - although confirming such a claim would require comparative evidence from a wider range of governance contexts.

### ***Research Gap***

Existing AI policy document analyses are geographically concentrated in Northern European and North American contexts (Toll et al., 2020; Persson et al., 2017). The Arab governmental context is absent from this comparative evidence base, leaving open the question of whether the structural value patterns observed in Scandinavian AI governance represent universal institutional logics or culturally specific phenomena. This study addresses that gap by applying a validated analytical framework to a systematically selected Saudi document corpus, enabling the first direct comparison between Saudi and Scandinavian AI governance value distributions.

## **Methodology**

### ***Research Design***

This study employs a mixed-method research design combining systematic qualitative content analysis with quantitative matrix coding (Hsieh & Shannon, 2005; Bazeley & Jackson, 2013). The qualitative layer captures nuanced textual meanings and contextual interpretations, while the quantitative layer produced via NVivo 15 matrix queries enables distributional comparison and cross-contextual benchmarking. This dual-layer approach responds to calls in the public values literature for methodological triangulation in governance document analysis.

### **Document Sample**

The sample comprises nine official documents produced by Saudi governmental entities between 2016 and 2024, selected through purposive sampling to ensure functional diversity including legislative, ethical, operational, and strategic document types:

This sampling logic aligns with Patton's (2015) concept of information-rich cases and Boddy's (2016) qualitative-saturation criterion: the corpus represents the complete available universe of Saudi AI governance documents in the period studied, satisfying conceptual sufficiency without claiming statistical representativeness.

### **Analytical Framework**

The Bannister & Connolly (2014) framework operationalises public values as three nested clusters. Duty-Oriented Values encompass legal compliance, institutional accountability, integrity, impartiality, and professionalism values derived from the Weberian conception of bureaucratic legitimacy. Service-Oriented Values include efficiency, effectiveness, transparency, and responsiveness values oriented toward optimal service delivery. Socially-Oriented Values encompass equity, fairness, privacy, and public accountability values concerned with collective welfare and citizen rights. This tripartite structure was mapped onto a NVivo node hierarchy and applied deductively to all nine documents.

### **NVivo Coding Procedure and Data Sources**

Three independent data matrices were generated in NVivo 15: (1) Document-level coding density (198 total references), measuring raw value salience per document; (2) Cross-tabulation matrix Q3 (222 diagonal references; 270 total including inter-category intersections), measuring pure category presence; and (3) Considerations-to-risks matrix Q4 (115 considerations, 54 risks;  $n = 169$ ), measuring rhetorical balance. These three matrices measure structurally different analytical dimensions and are not interchangeable. The diagonal references of Q3 were adopted as the primary ranking measure, as they capture the uncontaminated categorical identity of each coded segment, whereas row totals which inflate categories with higher inter-category overlap were retained as a secondary robustness check.

### **Trustworthiness and Rigour**

This study positions itself within a constructivist-interpretive paradigm of qualitative content analysis (Lincoln & Guba, 1985), in which validity is reframed as trustworthiness and the truth-claim of findings rests on the transparency, coherence, and auditability of the analytical procedure rather than on inter-rater statistical convergence. This paradigmatic anchoring is consequential: it determines which procedures count as evidentiary, which omissions constitute methodological gaps, and which alternative procedures, such as Cohen's Kappa coefficients, belong to a different epistemological tradition altogether (O'Connor & Joffe, 2020). The present section makes anchoring explicit and details the six interlocking mechanisms through which trustworthiness was operationalised in this study.

#### **1- The Four Criteria of Trustworthiness (Lincoln & Guba, 1985)**

Following the foundational framework of Lincoln and Guba (1985), four interlocking criteria guided the qualitative trustworthiness assessment of this study. Credibility was enhanced through prolonged

**Table 1. Document sample: Saudi Arabian AI policy corpus (2016–2024)**

Document	Year	Issuing Body
AI Adoption Framework	2023	SDAIA
AI Ethics Principles	2024	SDAIA
Generative AI for the Public Sector	2023	SDAIA
AI Implementation Guidelines	2022	SDAIA
SDAIA Institutional Publications	2016–24	SDAIA
Personal Data Protection Law Executive Regulations	2023	SDAIA/ NDMO
National AI Index	2023	Ministry of Economy
Personal Data Protection Law (PDPL)	2021	Royal Decree
Vision 2030	2016	Council of Eco- nomic Affairs

engagement with the documentary corpus, which was read in full prior to coding and re-engaged systematically across multiple coding cycles. Matrix coding queries in NVivo 15 generated transparent, traceable trails that allow any third-party reader to trace each interpretive claim back to its underlying textual reference. Patterns across the nine documents were triangulated against the Bannister and Connolly (2014) framework so that emerging interpretations were anchored in pre-established theoretical categories rather than ad hoc judgements.

Transferability is supported by thick description (Geertz, 1973) of the institutional and temporal context of the documents, the analytical framework, the coding procedures, and the corpus characteristics. The reader is therefore positioned to evaluate the applicability of the findings to other governmental contexts, whether Gulf, Arab, or non-Western, through informed comparison rather than unwarranted generalisation.

Dependability was addressed through the consistent application of a standardised coding framework (Bannister & Connolly, 2014) across all nine documents. NVivo's coding stripes, query outputs, and node-reference counts provide a stable audit trail in which every coding decision is preserved at the document-paragraph level, ensuring that the analytical process can be reconstructed in full by an external reviewer.

Confirmability is strengthened by grounding all interpretive claims in direct textual evidence from the policy documents, with NVivo reference numbers enabling independent verification of each coding decision against its source. Researcher-imposed bias is further constrained by the deductive nature of the coding scheme, in which categories were derived from the framework prior to any encounter with the corpus.

## 2- Audit Trail and Operational Transparency

Beyond the four core criteria, this study implements an explicit audit trail (Lincoln & Guba, 1985) within the NVivo project file. Every node, every coding stripe, every query, and every memo is preserved and retrievable. The NVivo project itself constitutes a primary methodological artefact: a third-party researcher could open the project, regenerate the matrix coding queries, and verify both the coding density figures (198 references; 222 pure-diagonal references in Q3; 115 considerations and 54 risks in Q4) and the underlying value attributions on a reference-by-reference basis. This level of operational transparency exceeds what is conventionally provided in qualitative reports and aligns with the decision-trail recommendation of Nowell, Norris, White and Moules (2017).

## 3- Thick Description of Coding Procedures

The coding procedure is documented in full in the methodology section. Each value category in the Bannister and Connolly (2014) framework was operationalised through definitional memos in NVivo before coding commenced; coding rules were refined iteratively in the first two documents, then locked for the remaining seven. This rule-based deductive procedure, distinct from emergent coding through open-ended thematic analysis, constitutes a fundamentally different epistemic operation: the analyst is not generating categories but applying a pre-specified theoretical lens to test its empirical traction in a non-Western governance corpus. This distinction is methodologically central to the trustworthiness argument that follows.

## 4- Reflexivity and Research Memos

Throughout the coding process, the researcher maintained reflexive memos within NVivo to document interpretive ambiguities, framework-adjacent edge cases, and the rationale for coding decisions in genuinely borderline references. These memos, attached to specific nodes and references, expose the researcher's analytical reasoning to external scrutiny and serve a dual purpose: documenting the cognitive process behind interpretive judgements and enabling the researcher to revisit and reconsider those judgements at later stages. Reflexivity is treated here not as confessional commentary but as a structural component of the audit trail.

## 5- Peer Debriefing

Peer debriefing was conducted at two checkpoints: after the first three documents had been coded, to ensure category-application consistency, and after all nine documents had been coded, to validate the matrix outputs and the dominance percentages reported in the results. The debriefing partner held a postgraduate qualification in public administration and was not otherwise involved in the project. Codings flagged as ambiguous were discussed and, where warranted, revised. This procedure satisfies the peer-debriefing recommendation of Lincoln and Guba (1985), and the corroborative-review function described by Nowell et al. (2017).

## 6- Rule-Based Consistency

The final mechanism is rule-based consistency in the sense developed by Elo and Kyngäs (2008) for qualitative content analysis and elaborated by Nowell et al. (2017) for thematic analysis. The Bannister and Connolly (2014) framework supplies the rules; these rules are applied uniformly across all references; deviations are documented in memos and audited at the end. Consistency, in this formulation, is achieved not through inter-coder convergence but through intra-procedural fidelity, namely the same analyst, the same rules, the same audit trail, applied without unauthorised drift.

## 7- On Inter-Coder Reliability and the Paradigmatic Boundaries of Cohen's Kappa

A reasonable reader may ask whether the absence of a second independent coder, and consequently the absence of Cohen's Kappa or Krippendorff's Alpha, constitutes a methodological limitation. The position taken here is that it does not, for reasons that are paradigmatic rather than apologetic.

Cohen's Kappa (Cohen, 1960; McHugh, 2012) and Krippendorff's Alpha (Krippendorff, 2018) emerge from a positivist tradition of content analysis that treats coding as a technical operation requiring inter-rater statistical convergence to validate the existence of an objective category in the text. Within that tradition, the analyst is essentially a measurement instrument; two such instruments must yield convergent readings for the measurement to be trusted. Landis and Koch (1977) provide the canonical thresholds ( $K \geq 0.61$  = substantial;  $K \geq 0.81$  = almost perfect), and McHugh (2012) refines the interpretive bands with attention to clinical-grade applications.

This study, however, does not operate within the positivist content-analysis tradition. It operates within the constructivist qualitative content-analysis tradition of Hsieh and Shannon (2005) and Elo and Kyngäs (2008), in which coding is theory-driven, deductive, and interpretive. In this tradition, the coding act is not a measurement, but an analytical reading guided by an explicit theoretical framework; reliability is reframed as dependability, and validity is reframed as credibility plus auditability.

O'Connor and Joffe (2020) review the inter-coder reliability literature and characterise the matter as an unsettled debate in which scholars disagree about whether Kappa-style coefficients should be required across all qualitative work or whether they belong specifically to positivist content analysis. Their conclusion, and the position adopted here, is that the appropriateness of inter-coder reliability statistics depends on the epistemological commitments of the study: required where the study presumes objective categories to be discovered, optional where the study deductively applies an explicit framework whose categories are theoretically pre-specified and whose consistency is established through audit trail and rule-based application.

This position is reinforced by Hammersley (2013), who emphasises that numerical indicators in qualitative analysis function as auxiliary evidence rather than conclusive proof, a stance fully consistent with the constructivist epistemology of this study.

The present study falls clearly into the latter case. The Bannister and Connolly (2014) framework supplies the categories; the analytical task is not to discover unknown categories but to test the framework's empirical reach on a non-Western corpus. The trustworthiness of that test rests on transparency and

traceability, both of which are fully delivered through the NVivo audit trail and the thick description above, rather than on a coefficient of inter-rater convergence.

It must therefore be acknowledged, for the avoidance of doubt, that a second-coder protocol would have produced a Kappa coefficient. It would not, however, have addressed any threat to validity that the present six-mechanism trustworthiness package does not already address. The decision against a second-coder protocol is methodologically deliberate, paradigmatically anchored, and aligned with the published recommendations of Lincoln and Guba (1985), Hsieh and Shannon (2005), Elo and Kyngäs (2008), Nowell et al. (2017), and O'Connor and Joffe (2020).

## Results

### Document-Level Coding Density

Table 2 presents the distribution of coding references across the nine documents ordered by decreasing density. Three documents collectively account for 63.6% of all references: the AI Adoption Framework (50 refs, 25.3%), the AI Ethics Principles (40 refs, 20.2%), and the Generative AI for the Public Sector document (36 refs, 18.2%). This concentration reflects the functional logic of value salience: operationally comprehensive documents that address the full governance lifecycle generate systematically denser value content than narrowly scoped legislative or strategic texts.

Two apparent anomalies warrant interpretive comment. Vision 2030 (n = 2, 1.0%) operates at the level of aspirational goals rather than procedural norms, generating few codeable value statements under the Bannister & Connolly framework. The Personal Data Protection Law (n = 5, 2.5%) despite its substantial citizen rights implications employs technical legal language that belongs to the domain of jurisprudence rather than public values discourse as operationalised in this framework. Neither observation diminishes their substantive importance; both reflect the limits of the coding instrument rather than the limits of the documents.

Document	Coding Density	N (%)
AI Adoption Framework		50 (25.3%)
AI Ethics Principles		40 (20.2%)
Generative AI for Public Sector		36 (18.2%)
AI Implementation Guidelines		22 (11.1%)
SDAIA Publications		19 (9.6%)
PDPL Executive Regulations		13 (6.6%)
National AI Index		11 (5.6%)
PDPL		5 (2.5%)
Vision 2030		2 (1.0%)

Source: NVivo 15 matrix coding output, 2026.

Figure 1. Document-level Coding Density Saudi AI Policy Documents (n = 198)

### Value Category Hierarchy Cross-Tabulation Matrix (Q3)

Table 3 presents the Q3 cross-tabulation matrix. Diagonal cells (shaded) represent pure coding references segments coded exclusively within a single category and constitute the primary ranking basis.

Table 4 summarises the category ranking by both metrics: It must be emphasised at the outset that the 43.7% figure is reported here as a descriptive quantitative observation about coding density distribution, not as a causal indicator. The interpretive remarks that follow concern patterns of value emphasis in the textual corpus rather than claims about the policy-maker's underlying intentions or the documents' real-world effects. Several analytical observations merit emphasis. First, Duty-Oriented Values dominate by a substantial margin (43.7%), consistent with the professionalism dominance found in European AI governance research (Toll et al., 2020; Persson et al., 2017). Second, the single-reference margin separating Socially-Oriented (63) from Service-Oriented (62) values signals a deliberately balanced policy discourse that attends to both citizen rights and service quality with near-equal weight. Third, the high internal cohesion scores particularly Socially-Oriented Values at 96.9% suggest that the Bannister & Connolly framework may possess meaning-

ful discriminative capacity in the Saudi context, with categories maintaining distinct conceptual identities. Fourth, the 22-reference cross-category intersection between Duty-Oriented and Service-Oriented Values the highest inter-category overlap in the matrix reflects the inherent conceptual proximity of transparency and accountability across these two value clusters.

### Implications of AI Integration for Public Values

#### 1- Implications for Duty-oriented Values

The AI Adoption Framework and PDPL Executive Regulations reveal a structurally novel legal liability challenge. Under Saudi data protection legislation, a data processor that exceeds the instructions of the data controller assumes full controller liability. In multi-layered AI environments where processing chains may involve primary contractors, sub-processors, and cloud infrastructure providers the operationalisation of this principle generates an acute boundary-definition problem: at what point does a processor's autonomous inference constitute deviation from controller instructions, thereby triggering a liability reclassification? This is not a hypothetical concern but an emerging operational challenge that existing Weberian accountability frameworks were not designed to address.

On the dimension of institutional responsibility, the AI Ethics Principles articulate a clear philosophical position:

*"It is essential to build and design a human-controlled AI system where decisions on the processes and functionality of the technology are monitored and executed and are susceptible to intervention from authorized users." (Saudi AI Ethics Principles, Reference 40)*

This formulation constitutes an implicit rejection of full algorithmic delegation. Yet it simultaneously generates an unresolved operational tension: how can effective human oversight of thousands of automated daily decisions be maintained without systematically negating the efficiency gains that motivate AI adoption? The documents assert the principle without resolving its practical operationalisation.

#### 2- Implications for Service-Oriented Values

The AI Ethics Principles specify a demanding operational standard for transparency:

*"data, algorithms, capabilities, processes, and purpose of the AI system need to be transparent and communicated as well as explainable to those who are directly and indirectly affected." (Saudi AI Ethics Principles, References 32–33)*

This standard confronts the well-documented accuracy-transparency trade-off in machine learning: as models increase in complexity and predictive accuracy, their internal decision pathways become

**Table 2. Document-level Coding Density Saudi AI Policy Documents (n = 198)**

Document	Total refs	% of total	Cumulative %
AI Adoption Framework	50	25.3%	25.3%
AI Ethics Principles	40	20.2%	45.5%
Generative AI for Public Sector	36	18.2%	63.6%
AI Implementation Guidelines	22	11.1%	74.7%
SDAIA Publications	19	9.6%	84.3%
PDPL Executive Regulations	13	6.6%	90.9%
National AI Index	11	5.6%	96.5%
PDPL	5	2.5%	99.0%
Vision 2030	2	1.0%	100.0%
Total	198	100%	

Source: NVivo 15, Matrix Coding Query, 2026.

**Table 3. Cross-tabulation Matrix (Q3) Inter-category Intersections**

Value Category (rows = coded categories)	Socially-Oriented (A)	Service-Oriented (B)	Duty-Oriented (C)
← Pure diagonal references per category	63	62	97
Duty-Oriented Values	1	22	97 ←
Service-Oriented Values	1	62 ←	22
Socially-Oriented Values	63 ←	1	1
Row totals	65	85	120

Source: NVivo 15, Q3 matrix coding query, 2026. ← denotes diagonal (pure) cell value.

**Table 4. Value category ranking by pure diagonal references and row totals (Q3)**

Rank	Value Category	Pure refs	% of 222	Row total	% of 270	Internal cohesion
1	Duty-Oriented Values	97	43.7%	120	44.4%	97/120 = 80.8%
2	Socially-Oriented Values	63	28.4%	65	24.1%	63/65 = 96.9%
3	Service-Oriented Values	62	27.9%	85	31.5%	62/85 = 72.9%
	Total (diagonal)	222	100%	270	100%	

Source: NVivo 15, Q3 matrix coding query, 2026.

progressively less interpretable even to their designers. Saudi documents assert full transparency as a non-negotiable principle without acknowledging this structural tension or specifying acceptable trade-off protocols. A similar productive ambiguity attends efficiency: the Generative AI document acknowledges that AI-generated content may carry “outdated, biased, or misleading information,” implying that quantitative efficiency gains (speed, volume) may not translate to qualitative efficiency gains (accuracy, trustworthiness).

### 3- Implications for Socially-Oriented Values

The AI Ethics Principles adopt a proactive, lifecycle-spanning anti-bias commitment:

*“The fairness principle requires taking necessary actions to eliminate bias, discrimination or stigmatization of individuals, communities, or groups in the design, data, development, deployment and use of AI systems.” (Saudi AI Ethics Principles, Reference 6)*

This formulation is analytically significant in two respects. First, it covers the complete AI development lifecycle from design to deployment rather than treating bias as a post-hoc correction problem. Second, it explicitly encompasses structural data bias (historical inequities embedded in training datasets) alongside intentional discrimination, acknowledging that algorithmic bias most frequently operates through apparently neutral quantitative processes rather than explicit discriminatory intent.

On public accountability, the documents establish audit requirements without specifying enforcement mechanisms who reviews audit reports, who investigates allegations of non-compliance, and who holds sanctioning authority. This gap between declared accountability structures and specified enforcement pathways represents a recurring pattern in AI governance documents globally (Jobin et al., 2019) and suggests that the institutional infrastructure for AI accountability is still under construction in the Saudi context.

### Challenges to Public Values Seven Structural Tensions

The NVivo qualitative analysis identified seven documented structural challenges to public values governance, each grounded in primary textual evidence (Table 5):

**Table 5. Structural Challenges to Public Values Document-grounded Evidence (n = 7 challenges)**

Challenge	Primary Textual Evidence	Value Category	Source
1 Strategic misalignment between AI initiatives and national priorities	“The most notable of these challenges include weak alignment between AI initiatives and national priorities, limited availability of suitable data for training, variations in the readiness of technical infrastructure across entities”	Duty-Oriented	National AI Index
2 Scarcity of specialised national AI talent	“A shortage of specialized national talent, and a lack of unified and clear measurement standards”	Duty-Oriented	National AI Index
3 Algorithmic bias and automated discrimination risk	“Automated decision-support technologies present major risks of bias and unwanted application at the deployment phase, so it is critical to set out mechanisms to prevent harmful and discriminatory results”	Socially-Oriented	AI Ethics Principles
4 Automated decision opacity the black-box problem	“This raises serious concerns about transparency and accountability, especially when such systems influence decisions that directly affect individuals or society”	Service-Oriented	AI Adoption Framework
5 Legal and ethical risks of generative AI in public service	“Adopting a generative AI system may carry legal and ethical implications... the risk of infringement of intellectual property rights, concerns about data protection, and the potential for human rights violations”	Service + Duty	AI Guidelines
6 Privacy erosion and personal data misuse	“Personal data is not misused and exploited, and the decision criteria of the automated technology is not based on personally identifying characteristics or information”	Socially-Oriented	AI Ethics Principles
7 Institutional trust deficit among staff and citizens	“Lack of clarity regarding the purpose of using AI in an organization and its connection to the needs of various stakeholders... may lead to losing their confidence in the organization”	Duty-Oriented	SDAIA Publications

Source: Qualitative content analysis in NVivo 15, 2026.

### Considerations-to-Risks Matrix (Q4) Quantifying Value Tensions

Table 6 presents the full Q4 matrix; Table 7 summarises the diagonal consideration-to-risk ratios per value category. The aggregate ratio of 2.13:1 (115 considerations : 54 risks) confirms that Saudi AI policy documents maintain an overall optimistic orientation, but one that is substantially more tempered than the 5.62:1 ratio reported for Swedish documents by Toll et al. (2020). The finding that risk acknowledgement accounts for 32% of total Q4 references (versus 9.6% in the Swedish corpus) is the study’s most consequential comparative result.

### NVivo Visual Outputs Treemap Analysis

Duty-Oriented Values 43.7%				Socially-Oriented Values 28.4%				Service-Oriented Values 27.9%				
Legal Compliance 30 refs (13.5%)	Institutional Accountability 19 refs	Integrity 16 refs	Professionalism 14 refs	Efficiency 10 refs	Transparency 8 refs	Public Accountability	Fairness & Equity	Data Privacy	Collective Protection	Service Transparency	Service Efficiency	Decision Quality

Source: NVivo 15 treemap output, 2026. Segment area proportional to pure reference count.

Figure 2. Treemap Value Category Distribution by Size and Sub-value Composition

The treemap (Figure 2) operationalises three analytical levels simultaneously, drawing on the macro/meso/micro analytical schema developed by Veale and Brass (2019) and operationalised within the AI-governance literature by Kuziemski and Misuraca (2020): the macro level is captured by the relative proportions among the three value categories; the meso level by the value subcategory distributions within each category; and the micro level by the Benefit/Consideration/Risk dimension within each subcategory.

Table 6. Considerations-to-risks Matrix (Q4) Full Inter-category Breakdown (n = 169)

Value Category	Consid.× Duty	Risk× Duty	Consid.× Service	Risk× Service	Consid.× Social	Risk× Social
Duty-Oriented	50 ◀	22 ◀	12	6	1	0
Service-Oriented	12	5	19 ◀	6 ◀	0	0
Socially-Oriented	1	0	0	0	20 ◀	15 ◀
Total	115 considerations 54 risks N = 169					

Source: NVivo 15, Q4 matrix coding query, 2026. ◀ = diagonal (within-category) cell.

Table 7. Diagonal Consideration-to-risk Ratios by Value Category (Q4 diagonal cells)

Value Category	Considerations	Risks	Total	Ratio (C: R)	Analytical interpretation
Duty-Oriented	50	22	72	2.27 : 1	Highest absolute risk count
Service-Oriented	19	6	25	3.17 : 1	Most optimistic category
Socially-Oriented	20	15	35	1.33 : 1	◀ Highest relative tension
All categories (aggregate)	115	54	169	2.13 : 1	Vs. 5.62:1 in Sweden

Source: NVivo 15 diagonal cells, Q4 matrix, 2026.

Three observations emerge. First, the proportional dominance of the Duty-Oriented segment provides visual confirmation of the 43.7% findings. Second, Legal Compliance and Institutional Accountability appear across all three rhetorical dimensions (Benefit, Consideration, Risk), confirming that Saudi documents engage with these values through a multidimensional rather than uniformly promotional lens. Third, within the Socially-Oriented segment, Accountability and Fairness show more evenly distributed Benefit-Consideration-Risk profiles than their counterparts in other categories a visual correlate of the category’s 1.33:1 tension ratio.

Word frequency analysis confirms the rhetorical pattern at the lexical level. The most frequent terms data (1,185 occurrences), personal (390), controller (196), processing (189) form a legal-compliance semantic cluster dominated by Duty-Oriented vocabulary. The prevalence of ensure (220 occurrences) as a high-frequency modal verb exemplifies the mandatory rhetorical register that distinguishes Saudi documents from the aspirational-permissive register (can/may/will be able to) that Toll et al. (2020) identified as characteristic of Swedish AI policy discourse.

## Discussion

### *Cross-Contextual Convergence The 43.7% Structural Pattern*

The cross-contextual reading offered below is bounded by a deliberate methodological caution against cultural essentialism (Said, 1978; Mahmood, 2005). The patterns identified are interpreted as institutional rather than civilisational; that is, the convergence between Saudi and Swedish policy emphases is treated as a finding about how governments-as-institutions frame transformative technologies under conditions of accountability pressure, not as evidence of a stable civilisational trait. Cultural elements enter the interpretation only as one among several institutional vectors, and emphatically not as essentialised dispositions of any society or political tradition.

This anti-essentialist stance is further grounded in Asad (2003), who argues that culture is a process of practices and discourses formed within fields of power and knowledge rather than a fixed essence; the convergence between Saudi and Swedish patterns is therefore read as an institutional rather than a civilisational phenomenon.

The precise numerical correspondence between Saudi Duty-Oriented Values ( $97/222 = 43.7\%$ ) and Swedish professionalism values ( $228/522 = 43.7\%$ ) represents the study's most analytically provocative finding. Two explanatory mechanisms are proposed. The conceptual equivalence mechanism: "professionalism" in the Rose et al. (2015) framework encompasses legitimacy, robustness, accountability, and infrastructure the same substantive content measured by "Duty-Oriented Values" in Bannister & Connolly (2014). The mathematical convergence therefore reflects genuine conceptual correspondence rather than coincidental numerical agreement. The structural governance mechanism: state institutions, by their constitutive definition, derive legitimacy from legal authority and public accountability. Accordingly, when governments encounter potentially disruptive technology, the institutional reflex is to subordinate its transformative potential to existing frameworks of legality and control. This appears to reflect structural features of governmental institutions more than culturally specific responses — an interpretation consistent with comparable observations in Saudi, Swedish, and Danish (Persson et al., 2017) governance contexts.

Key finding: The 43.7% convergence between Saudi and Swedish AI governance documents across different analytical frameworks, document sets, time periods, and civilisational contexts contributes to emerging evidence for a possible structural pattern in governmental technology governance.

### *Cross-Contextual Divergence the Social Values Gap*

The most substantial point of divergence between the two contexts is the Socially-Oriented Values gap: 28.4% in Saudi Arabia versus 7.5% (engagement values) in Sweden a differential of 20.9 percentage points. Two causal mechanisms are proposed. The document composition mechanism: the Saudi sample includes PDPL and its executive regulations comprehensive data protection legislation — which by their regulatory logic generate high Socially-Oriented value density. The Swedish sample contains no equivalent legislative document. The civilisational normative mechanism: Saudi governance traditions conceptualises justice, privacy, and collective protection as divine obligations rather than policy preferences. This invests Socially-Oriented values with a normative force that may generate higher salience in governance discourse independent of regulatory requirements. Disentangling these two mechanisms document composition versus civilisational normative pressure requires a controlled comparative design that matches document types across contexts, a task for future research.

### *Temporal Maturation A Second-Generation Governance Model*

The 2.13:1 aggregate consideration-to-risk ratio of Saudi documents, compared to the 5.62:1 benefits-to-risks ratio of Swedish documents (Toll et al., 2020), is partially attributable to the temporal distance between the two study periods. Between 2017–2018 (when the Swedish documents were produced) and 2016–2024 (the Saudi sample period), the global AI governance environment experienced a fundamental shift: the algorithmic bias scandals affecting criminal justice and credit scoring systems became widely publicised; GDPR entered into force, establishing a binding data rights precedent; the EU AI Ethics Guide-

lines introduced a comprehensive value framework; and academic critique of AI optimism intensified. Saudi policymakers in 2016–2024 accordingly operated within a knowledge environment that had already processed these lessons. The higher risk acknowledgement ratio in Saudi documents therefore partly reflects the advantage of temporal distance rather than exclusively the influence of institutional or civilisational factors a finding that should appropriately qualify optimistic readings of Saudi AI governance maturity.

### *The Absent Dimension Participatory Values*

Both the Swedish and Saudi corpora are characterised by the near-absence of engagement and participatory values 7.5% in Sweden, and a similarly marginal presence in Saudi Arabia despite explicit commitments to public transparency and accountability. This convergence is consistent with the pattern observed by Chadwick & May (2003) across multiple e-government contexts: participation ideals tend to receive less emphasis than efficiency and compliance priorities in governmental technology discourse. The finding raises a methodological question about the limits of the Bannister & Connolly framework itself: if participatory values are structurally absent from governmental AI documents globally, their absence may reflect the limitations of the coding instrument designed for democratic governance settings rather than a deficit in the governance systems under analysis.

## Comparative Synthesis : Saudi Arabia and Sweden

Table 8 presents the comprehensive cross-contextual comparison:

**Table 8. Comprehensive Cross-contextual Comparison: Saudi Arabia vs. Sweden**

Dimension	Saudi Arabia (2016–2024)	Sweden 2017–2018 (Toll et al., 2020)
Analytical Framework	Bannister & Connolly (2014): 3 categories	Rose et al. (2015): 4 value ideals
Sample	9 documents 198 refs 222 pure diagonal (Q3)	10 documents 522 coded statements
Leading value category	Duty-Oriented: 97/222 = 43.7%	Professionalism: 228/522 = 43.7%
← Convergence note	Genuine conceptual equivalence: both categories measure legality, accountability, institutional integrity	Supports universal structural governance logic confirmed also in Denmark (Persson et al., 2017)
Second value category	Socially-Oriented: 28.4% [+20.9 pts]	Efficiency: 30.1%
Weakest category	(framework has 3 categories only)	Engagement: 7.5%
Risk acknowledgement	54/169 = 32.0%	50/522 = 9.6%
Consideration: Risk ratio (aggregate)	2.13 : 1	5.62 : 1 (benefits: risks) “overly optimistic”
Dominant rhetorical register	Mandatory: ensure / shall / must (220+ occurrences)	Aspirational: can / may / will be able to
← Key comparative finding	Saudi documents represent the “more nuanced view” called for by Toll et al. acknowledging risks at 32% vs. 9.6%	

Source: Present Study (NVivo 15, 2026) and Toll et al. (2020, Figure 1: 53).

## Conclusions

### *Summary of Key Findings*

- **Finding 1 Cross-contextual convergence in dominance of duty/professionalism values:** Duty-Oriented Values dominate Saudi AI policy discourse at 43.7%, precisely mirroring the professionalism dominance found in Swedish AI governance (Toll et al., 2020) and Danish e-government strategies (Persson et al., 2017). This tri-contextual convergence suggests a possible institutional pattern in which governments frame transformative technologies primarily through compliance and control lenses, across the governance contexts examined.
- **Finding 2 Near-equality of Socially-Oriented and Service-Oriented values:** The single-reference margin separating Socially-Oriented (63) from Service-Oriented (62) values indicates a deliberately balanced policy discourse an unusual degree of attentiveness to both citizen rights and service quality that may reflect the dual normative pressures of national governance norms and Vision 2030 service improvement targets.

- **Finding 3 Socially-Oriented Values as the highest-tension category:** The Q4 matrix reveals that Socially-Oriented Values carry the most acute value tensions (1.33:1 consideration-to-risk ratio), with risks approaching parity with considerations a finding with direct policy implications for AI fairness, privacy, and accountability governance.
- **Finding 4 Second-generation governance model:** Saudi AI policy documents represent what might be termed a second-generation governance model: one that moves beyond first-generation promotional optimism toward legislative caution, acknowledging risks at 32% of Q4 references versus 9.6% in the Swedish corpus. This embodies precisely the “more nuanced view” advocated by Toll et al. (2020).

### *Theoretical Contributions*

This study makes three contributions to public values and AI governance literature. First, it provides the first empirical evidence from an Arab governmental context suggesting that duty/professionalism value dominance may extend beyond Scandinavian governance contexts in AI governance, extending the geographic scope of prior findings beyond Scandinavia. Second, it proposes and operationalises the distinction between “pure diagonal” and “row total” measures in NVivo matrix coding as a methodological refinement with substantive analytical consequences for category ranking. Third, it introduces the concept of a second-generation AI governance model characterised by higher risk acknowledgement and mandatory rhetorical registers as a temporal framework for interpreting cross-contextual governance comparisons.

### *Limitations and Future Research*

The interpretive scope of this study should be qualified by an explicit distinction between two modes of generalisation. Statistical generalisation, the inferential extension of sample-based findings to a defined population, is neither claimed nor sought; the corpus of nine documents was not assembled as a probabilistic sample. What is claimed is analytical generalisation in the sense developed by Yin (2018) for purposive qualitative inquiry: the empirical fit between a non-Western governance corpus and the Bannister and Connolly (2014) framework, together with the cross-corpus correspondence to Toll et al. (2020), supports the framework’s conceptual reach into a new institutional context. Future research employing larger purposive samples, or comparative samples across additional Gulf and Arab governments, would extend the analytical reach without necessarily transforming the design into a statistically generalisable one.

Three limitations bound the study’s interpretive scope. The document sample ( $n = 9$ ), while representing the full universe of available Saudi AI governance documents in the study period, limits statistical generalisation. The cross-framework comparison with Toll et al. (2020) requires cautious interpretation given the different analytical frameworks applied. And the 2016-2024 study period means that documents produced under Saudi Arabia’s more recent AI governance initiatives fall outside the analytical scope.

Future research should pursue three avenues:

- 1- Controlled document-matched comparisons across diverse institutional and cultural AI governance contexts, holding document functional type constant;
- 2- Longitudinal analysis of Saudi AI governance documents to track value evolution as the governance regime matures;
- 3- Qualitative interview-based research with Saudi AI policymakers to investigate the institutional logics that produce the observed value distributions.

## References

- Bannister, F., & Connolly, R. (2014). ICT, Public values and transformative government: A framework and programme for research. *Government Information Quarterly*, 31 (1), 119-128. <https://doi.org/10.1016/j.giq.2013.06.002>
- Asad, T. (2003). *Formations of the Secular: Christianity, Islam, Modernity*. Stanford University Press.
- Bazeley, P., & Jackson, K. (2013). *Qualitative Data Analysis with NVivo*. (2nd ed.). SAGE Publications.
- Boddy, C. R. (2016). Sample size for qualitative research. *Qualitative Market Research: An International Journal*, 19 (4), 426-432. <https://doi.org/10.1108/QMR-06-2016-0053>
- Bozeman, B. (2007). *Public values and public interest: Counterbalancing economic individualism*. Georgetown University Press.
- Bullock, J. B. (2019). Artificial intelligence, discretion, and bureaucracy. *The American Review of Public Administration*, 49 (7), 751-761. <https://doi.org/10.1177/0275074019856123>
- Chadwick, A., & May, C. (2003). Interaction between states and citizens in the age of the Internet: "e-government" in the United States, Britain, and the European Union. *Governance*, 16 (2), 271-300. <https://doi.org/10.1111/1468-0491.00216>
- Chen, Y.-C., Ahn, M. J., & Wang, Y.-F. (2023). Artificial intelligence and public values: Value impacts and governance in the public sector. *Sustainability*, 15 (6), 4796. <https://doi.org/10.3390/su15064796>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20 (1), 37-46. <https://doi.org/10.1177/001316446002000104>
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62 (1), 107-115. <https://doi.org/10.1111/j.1365-2648.2007.04569.x>
- Geertz, C. (1973). *The Interpretation of Cultures: Selected Essays*. Basic Books.
- Hammersley, M. (2013). *What is qualitative research?* Bloomsbury Academic.
- Hood, C. (1991). A public management for all seasons? *Public Administration*, 69 (1), 3-19. <https://doi.org/10.1111/j.1467-9299.1991.tb00779.x>
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15 (9), 1277-1288. <https://doi.org/10.1177/1049732305276687>
- Janowski, T. (2015). Digital government evolution: From transformation to contextualisation. *Government Information Quarterly*, 32 (3), 221-236. <https://doi.org/10.1016/j.giq.2015.07.001>
- Jobin, A., Lenca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1 (9), 389-399. <https://doi.org/10.1038/s42256-019-0088-2>
- Jæger, B., & Löfgren, K. (2010). The history of the future: Changes in Danish e-government strategies 1994–2010. *Information Polity*, 15 (4), 253-269. <https://doi.org/10.3233/IP-2010-0217>
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. (4th ed.). SAGE Publications.
- Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic countries. *Telecommunications Policy*, 44 (6), 101976. <https://doi.org/10.1016/j.telpol.2020.101976>

- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33 (1), 159-174. <https://doi.org/10.2307/2529310>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. SAGE Publications.
- Mahmood, S. (2005). *Politics of Piety: The Islamic Revival and the Feminist Subject*. Princeton University Press.
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22 (3), 276-282. <https://doi.org/10.11613/BM.2012.031>
- Ministry of Economy and Planning. (2023). *National AI Index*. Kingdom of Saudi Arabia.
- Moore, M. H. (1995). *Creating Public Value: Strategic Management in Government*. Harvard University Press.
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16 (1), 1-13. <https://doi.org/10.1177/1609406917733847>
- Patton, M. Q. (2015). *Qualitative Research and Evaluation Methods*. (4th ed.). SAGE Publications.
- Pollitt, C., & Bouckaert, G. (2017). *Public Management Reform: A Comparative Analysis*. (4<sup>th</sup> ed.). Oxford University Press.
- O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, 19, 1-13. <https://doi.org/10.1177/1609406919899220>
- Persson, J., Reinwald, A., Skorve, E., & Nielsen, P. (2017). Value positions in e-government strategies: Something is (not) changing in the state of Denmark. *ECIS 2017 Proceedings*, 904-917.
- Rose, J., Persson, J. S., Heeager, L. T., & Irani, Z. (2015). Managing e-government: Value positions and relationships. *Information Systems Journal*, 25(5), 531-571. <https://doi.org/10.1111/isj.12052>
- Said, E. W. (1978). *Orientalism*. Pantheon Books.
- Saudi Council of Economic Affairs and Development. (2016). *Vision 2030*. Kingdom of Saudi Arabia.
- SDAIA. (2023). *Generative AI Guide for Government Entities*. Saudi Data & AI Authority.
- SDAIA/NDMO. (2023). *Personal Data Protection Law Executive Regulations*. Kingdom of Saudi Arabia.
- SDAIA. (2024a). *AI Adoption Framework*. Saudi Data & AI Authority.
- SDAIA. (2024b). *AI Ethics Principles*. Saudi Data & AI Authority.
- Toll, D., Lindgren, I., Melin, U., & Madsen, C. O. (2020). Values, benefits, considerations and risks of AI in government: A study of AI policy documents in Sweden. *Journal of eDemocracy and Open Government*, 12(1), 40–60. <https://doi.org/10.29379/jedem.v12i1.593>
- Veale, M., & Brass, I. (2019). Administration by algorithm? Public management meets public sector machine learning. In K. Yeung & M. Lodge (Eds.), *Algorithmic Regulation*, (pp. 121-149). Oxford University Press.
- Yin, R. K. (2018). *Case study Research and Applications: Design and Methods*. (6<sup>th</sup> ed.). SAGE Publications.

