Japan Bilingual Publishing Co.

Philosophy and Realistic Reflection

https://ojs.bilpub.com/index.php/prr

ARTICLE

Delegating Authority to Algorithms: Legitimacy and Public Values in European Tax Administrations

Pascal Stiefenhofer [1*], *Ufuk Gergerlioğlu* [2], *Cafer Deniz* [1,3]

[1] *Department of Economics, Newcastle University, Newcastle upon Tyne NE1 4SE, UK*

[2] *Department of Economics, Hitit University, Çorum 19169, Türkiye*

[3] *Department of Economics, The University of Uşak, Uşak 64200, Türkiye*

ABSTRACT

This paper develops a philosophical and empirical framework for evaluating the democratic legitimacy of algorithmic decision-making in European tax administrations. Drawing on evidence from fifteen jurisdictions, it investigates how four core public values—intelligibility, contestability, non-domination, and equality before the law—are articulated in policy discourse and translated into institutional practice within AI-enabled tax governance. Using a large corpus of legal texts, policy strategies, and oversight reports, the study codes qualitative insights into a structured binary feature matrix and analyses this dataset using Multiple Correspondence Analysis and hierarchical clustering. The results reveal four recurrent constellations of algorithmic governance: pragmatic enforcement, rights-anchored governance, constitutional caution, and data-intensive profiling, with Finland exhibiting a distinctive pattern of constitutional and procedural restraint. These empirical configurations form the foundation for the paper's central philosophical question: Under what conditions can discretionary authority be legitimately delegated to algorithms? The analysis shows that although European administrations routinely appeal to ethical principles and human-centred AI rhetoric, many implementations transform these principles into managerial slogans, thereby weakening procedural justice, diminishing intelligibility, and limiting avenues for contestation. The paper argues that algorithmic decision-making can be democratically legitimate only when it sustains the four public-value conditions that anchor administrative authority. It concludes by proposing a reconceptualization of algorithmic

*CORRESPONDING AUTHOR:
Pascal Stiefenhofer, Department of Economics, Newcastle University, Newcastle upon Tyne NE1 4SE, UK;
Email: pascal.stiefenhofer@newcastle.ac.uk

ARTICLE INFO
Received: 9 September 2025 | Revised: 4 November 2025 | Accepted: 11 November 2025 | Published Online: 19 November 2025
DOI: https://doi.org/10.55121/prr.v3i1.757

CITATION
Stiefenhofer, P., Gergerlioğlu, U., Deniz, C., 2026. Delegating Authority to Algorithms: Legitimacy and Public Values in European Tax Administrations. Philosophy and Realistic Reflection. 3(1): 1–21. DOI: https://doi.org/10.55121/prr.v3i1.757

legitimacy as the preservation of intelligibility, contestability, non-domination, and equality within computationally mediated public administration.

Keywords: Artificial Intelligence; Tax Administration; Tax Audit; Ethics

1. Introduction

Artificial intelligence (AI) is not merely a technical instrument but a normative infrastructure that reorganizes institutional power, reshapes administrative discretion, and tests the ethical foundations of public governance [1, 2]. This transformation is particularly pronounced in tax administration, where the state's coercive authority intersects with principles of legality, proportionality, and equality before the law. The automation of fiscal functions—risk-based auditing, predictive analytics, and fraud detection—extends administrative capacity yet simultaneously reconstitutes the moral architecture of decision-making. By determining who is subjected to scrutiny, on what evidentiary grounds, and through which channels of contestation, algorithmic systems mediate relations of power and responsibility within the state [3, 4].

Within public administration theory, a persistent tension between efficiency and legitimacy has become newly visible in the age of algorithmic governance [5, 6]. As algorithmic systems increasingly determine outcomes that shape citizens' legal and financial standing, foundational democratic values—intelligibility, contestability, privacy, and justice—are placed under strain. Unlike traditional bureaucratic discretion, AI-driven decision-making often lacks identifiable agents of responsibility and operates through opaque computational processes, the so-called "black box", thereby undermining norms of reason-giving and answerability that sustain administrative legitimacy [7–11]. The challenge is therefore not merely technical but philosophical. If discretionary public authority is transferred to algorithms, what conditions must be met for such delegation to remain democratically legitimate?

This paper addresses that question directly. It asks: under what conditions can the delegation of discretionary authority to algorithmic systems in tax administration be compatible with democratic legitimacy, and how do contemporary European practices measure against those conditions?

We pursue this question by formulating legitimacy requirements grounded in public reason and procedural justice. These requirements set out four jointly necessary conditions for the democratic use of algorithmic authority. The first is intelligibility, which concerns the capacity to understand and reconstruct the reasons that guide a decision. The second is contestability, referring to the existence of effective opportunities for challenge and redress. The third is non-domination, which protects individuals from opaque or arbitrary control. The fourth is equality before the law, ensuring that comparable cases are treated consistently and without unjustified discrimination. We then examine how European tax administrations meet or fail to meet these conditions in practice.

The European context provides an instructive field of inquiry. The European Union promotes a strong normative vocabulary through instruments such as the General Data Protection Regulation (GDPR) and the proposed AI Act, yet national practices diverge sharply. They range from approaches centred on human oversight and public accountability to those that prioritise efficiency and deterrence [12, 13]. To examine these variations systematically, we combine comparative philosophical analysis with an empirical mapping of administrative practice across fifteen European tax administrations. Qualitative evidence drawn from legislation, policy documents, and oversight reports is converted into a binary feature matrix described in Section 4 and examined through Multiple Correspondence Analysis (MCA) and hierarchical clustering. These complementary methods reveal coherent families of practice that would otherwise remain obscured in narrative discussion and provide an evidential foundation for the normative evaluation developed in the following sections.

The findings presented in this study are twofold. The first concerns the diversity of national approaches. Although governments across Europe appeal to a shared rhetoric of human-centred artificial intelligence, they do not converge on a single model of governance. The machine-learning analysis identifies four distinct configurations of practice—pragmatic enforcement, rights-anchored governance, constitutional caution, and data-intensive profiling—with Finland standing as a structural outlier. The second finding concerns the normative implications of these patterns. Only some configurations come close to meeting the legitimacy conditions defended in this paper. Where intelligibility and contestability are weak, as in cases of proprietary opacity or cross-database profiling without effective redress, the moral justification for algorithmic authority becomes fragile, even when efficiency gains appear substantial.

By relating these results to current debates on public reason, digital power, and administrative justice, the paper develops a philosophical understanding of algorithmic legitimacy that is informed by comparative evidence. Its contribution rests on three elements. It offers a clarified normative standard for the delegation of authority to AI in taxation. It provides a cross-national mapping of European administrative practices in relation to that standard. It also proposes a methodological bridge that connects interpretive philosophy with evidence drawn from machine-learning-supported comparison. The remainder of the paper elaborates the conceptual framework in Sections 2 and 3, presents the comparative analysis in Section 4, and explores the ethical implications and normative tensions in Section 5 before concluding with implications for both democratic theory and governance design in Section 6. Two appendices, **Appendix A.1** and **Appendix A.2**, provide additional methodological details and supporting materials.

2. Theoretical Framework: Ethics, Discretion, and the Conditions of Legitimate Authority in AI-Governed Tax Administration

The incorporation of artificial-intelligence systems into public administration—most visibly within taxation—demands more than an assessment of their technical efficiency. It requires an inquiry into the normative conditions under which administrative authority continues to be legitimate. Historically, tax administration has stood at the intersection of bureaucratic rationality and legal obligation, balancing procedural efficiency with the moral demand for justification. The current migration from embodied discretion to computational decision-making thus represents not a mere technological refinement but a reconfiguration of judgment, accountability, and legitimacy within the modern state.

Legitimate authority in public administration depends on the presence of identifiable agents capable of reasoning, offering justification, and assuming responsibility for their actions. Authority is not reducible to the capacity to decide; it is constituted through *reason-giving* within a shared normative order. The delegation of decision-making to AI systems transforms this structure. Where human officials deliberate, AI systems infer—producing outcomes through probabilistic computation rather than through intentional reasoning or moral judgment. Such systems can optimise prediction and consistency, yet they cannot themselves satisfy the justificatory conditions that ground legitimacy. What emerges is a displacement of discretion from the sphere of deliberative reasoning to that of procedural inference, accompanied by a diffusion of accountability across technical and institutional domains. The central question, therefore, is how legitimacy can be sustained when the capacity for judgment and justification is externalised to computational architectures that act without intention or understanding.

To approach this question, it is first necessary to clarify what *artificial intelligence* denotes in this study. In the field of public administration, the term has acquired a specific institutional meaning. The OECD defines an AI system as *a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments* [13]. Wirtz, Weyerer, and Geyer [1] describe AI in the public sector as the application of data-driven, rule-based, or learning systems to automate or augment administrative decision-making. Bullock [3] emphasises that such systems reconfigure bureaucratic discretion and form, while Zouridis, van Eck, and Bovens [5] show that discretion itself becomes embedded in software code. Yeung [2] interprets AI governance as a *socio-technical assemblage* through which algorithmic rules mediate power and normativity, and Veale and Brass [14] identify this as a shift from explicit rule-making to *administration by algorithm*. Taken together, these perspectives indicate that AI in public administration must be understood not simply as a technology but as a decision infrastructure that transforms the exercise of administrative authority.

To situate this conception historically, the idea of "intelligent" machines has undergone successive transformations, each redefining what counts as reasoning and, by implication, what kind of agency a system may be said to exercise.

The Symbolic or Classical Era (1950s–1980s): Early AI, or *Good-Old-Fashioned AI* (GOFAI), conceived intelligence as symbolic rationalism—the manipulation of explicit rules and representations to emulate human reasoning. Following Turing [15] and McCarthy [16], intelligence was equated with formal logic and rule-based problem solving. This paradigm, codified by Russell and Norvig [17], aimed to build systems that *think and act rationally*. In administrative settings, such systems could, in principle, formalise decision procedures but lacked the capacity for learning or moral interpretation.

The Connectionist and Behavioural Turn (1980s–2000s): The brittleness of symbolic systems prompted a shift toward connectionist and behavioural models of intelligence. McClelland and Rumelhart [18] conceptualised cognition as distributed activation in neural networks, while Brooks [19] developed behaviour-based robotics in which intelligence emerged through adaptive interaction with the environment. As Cao [20] notes, this marked a movement *from object intelligence to system intelligence*: AI became dynamic, contextual, and behaviourally grounded.

The Data-Driven and Deep-Learning Phase (2000s–2015): With the rise of high-performance computing and vast data sets, AI evolved into a discipline of statistical optimisation. LeCun, Bengio, and Hinton [21] described deep learning as hierarchical feature extraction from massive data, enabling predictive and classificatory capacities at unprecedented scale. Intelligence was now *inferred from data rather than encoded in logic*, extending administrative applications while introducing new forms of opacity and bias.

The Reflective or Metasynthetic Age (2015–present): Contemporary AI, described by Cao [20] as *reflective* or *metasynthetic*, reconceives intelligence as an ecosystem of interacting human, social, and computational processes. It integrates perception, reasoning, learning, action, and reflection across technical and institutional layers. Mitchell [22] and Kissinger, Schmidt, and Huttenlocher [23] emphasise that the frontier of AI lies not in autonomy but in human integration—designing systems that are transparent, accountable, and sustainable. Intelligence in this sense is not a property of machines alone but of socio-technical systems embedded in legal, ethical, and organisational contexts.

Synthesising these traditions yields a definition suited to the present inquiry: *Artificial intelligence, in this study, denotes a class of computational architectures that integrate symbolic reasoning, statistical learning, and adaptive feedback to perform perception, inference, and decision-making within socio-legal environments. In the context of public administration, AI systems function as distributed decision infrastructures that extend administrative reasoning beyond the human agent while encoding normative assumptions within data, models, and design choices.*

This derivation clarifies that the "intelligence" of AI is neither singular nor autonomous: it is a hybrid of algorithmic inference and institutional framing. Within tax administration, AI manifests through predictive analytics, risk-scoring algorithms, and cross-database profiling systems. These instruments are not value-neutral tools but normatively charged mechanisms that implicitly define what counts as compliance, suspicion, or fairness. Assessing their legitimacy, therefore, requires tracing how computational structures interact with the moral architecture of public authority and determining whether the delegation of discretion to non-intentional systems can satisfy the justificatory conditions of legitimate rule.

2.1. From Bureaucratic Rationality to Algorithmic Governance

In classical administrative theory, legitimate authority rests on the rational–legal order (Weber [24]), which reconciles instrumental efficiency with rule-governed fairness. Bureaucracy, in this sense, derives its moral force from the *predictability of rule* and the *accountability of office*. The emerging "algorithmic turn" unsettles this equilibrium. As Zouridis, van Eck, and Bovens [5] demonstrate, the discretionary judgment once exercised by *street-level bureaucrats*—agents capable of context-sensitive interpretation—is being supplanted by *system-level bureaucracies* in which normative reasoning is embedded in software code. Decision-making, once dialogical and justificatory, is recast as statistical prediction. In these architectures, agency becomes distributed across designers, administrators, and artefacts, rendering the locus of responsibility opaque.

The ethical implications are profound. Danaher [25] identifies in such systems a *responsibility gap*: the diffusion of causal efficacy across human and non-human components erodes the moral link between action and accountability. The logic of this dilemma can be expressed biconditionally. *If* *and only if* responsibility presupposes an identifiable agent capable of intentional action, and *if* algorithms lack such agency, *then* responsibility cannot meaningfully attach to them. Conversely, when decision-making authority is delegated to algorithmic processes, the performative subject of public power—the official who can reason, explain, and be held to account—effectively disappears. What remains is an apparatus that performs the functions of authority without its moral predicates. The ethical coherence of bureaucratic legitimacy is thereby placed in question: rule by code risks becoming authority without agency, governance without responsibility.

2.2. The Four Normative Conditions of Legitimate Authority

To evaluate legitimacy under conditions of algorithmic governance, four interdependent principles are specified as jointly sufficient and individually necessary**:** *intelligibility*, *contestability*, *privacy*, and *fairness*. Authority delegated to artificial systems retains its normative status *if and only if* these four conditions are simultaneously satisfied.

1. Intelligibility: Derived from Fuller's [26] account of the internal morality of law and Habermas's [27] discourse theory of legitimacy, intelligibility denotes the capacity of administrative acts to appear as *reason-giving* within a shared normative order. Following Raz [28], authority is legitimate only when its directives can be understood as providing reasons for compliance. The epistemic opacity of machine-learning models therefore violates the condition of intelligibility and, by entailment, undermines legitimacy: where reasoning becomes inscrutable, obligation loses its moral foundation.
2. Contestability: Grounded in republican and procedural theories of non-domination, contestability requires that citizens possess both the *right* and the *practical capacity* to challenge administrative decisions [29, 30]. Authority is legitimate *if* those subject to it can demand justification and obtain correction through institutional means. Where algorithmic systems foreclose such review, governance devolves into domination—a relation of power without recourse.
3. Privacy: Extending Kant's injunction to treat persons as ends into the informational domain, privacy secures the moral and epistemic space within which autonomy is exercised [31, 32]. Continuous data extraction and surveillance negate this space. Hence, *if* autonomy presupposes privacy, and *if* algorithmic surveillance negates privacy, *then* such systems necessarily compromise autonomy and, derivatively, legitimacy. Privacy is thus not a contingent preference but a constitutive condition of moral agency.
4. Fairness: Building on Rawls [33] and later accounts of structural injustice, fairness demands that administrative systems allocate burdens and benefits in ways consistent with equality. As Binns [34] argues, the appearance of algorithmic neutrality is often a moral fiction: fairness, therefore, requires the active correction of bias. Formally, *if and only if* outcomes are non-arbitrary with respect to morally relevant differences does administrative power count as just.

Taken together, these principles define the biconditional of legitimate delegation:

$$\text{Delegation of discretion to AI is legitimate} \Leftrightarrow (\text{Intelligibility} \wedge \text{Contestability} \wedge \text{Privacy} \wedge \text{Fairness}).$$

Each principle is necessary—the violation of any one invalidates legitimacy—and their conjunction is sufficient, providing the complete set of conditions under which authority delegated to AI can remain normatively justified within a democratic order.

This biconditional serves both as a normative constraint and as an evaluative instrument. It constrains the moral permissibility of technological delegation by defining what must hold for legitimacy to persist, and it provides a logical structure for empirical assessment. Section 3 operationalises these four conditions into measurable dimensions, specifying the indicators, coding procedures, and comparative methods used to test whether the observed practices of fifteen European tax administrations satisfy the normative conjunction. In this way, the argument established here performs a dual function: it articulates the philosophical grammar of legitimate authority and simultaneously defines the hypothesis structure for its empirical validation.

2.3. Authority as a Performative and Institutional Phenomenon

Administrative authority, as employed in this analysis, is understood through the lens of Speech-Act Theory and the ontology of institutional facts. In his foundational account, Austin [35] demonstrated that authoritative acts do not merely describe obligations but *create* them through performative utterances: to declare a tax liability, to issue a penalty, or to authorise an audit is to bring a new institutional state of affairs into existence. Building on this, Searle [36, 37] conceptualises authority as a *status function* sustained by collectively recognised constitutive rules—the grammar of the institutional world in which "X counts as Y in context C". Within such a framework, an act becomes binding only when its conditions of satisfaction are fulfilled and its declaration is recognised as valid within the relevant institutional order.

Vanderveken [38] deepens this analysis by specifying the *direction of fit* that confers normative force upon an illocution: words shape the world when their performance, under appropriate conventions, enacts the reality to which they refer. It is through this performative alignment that administrative acts acquire legitimacy; they are not empirical descriptions but *normative creations* whose authority depends on rule-governed performance within a community of recognition.

Delegating discretionary power to algorithmic systems destabilises this performative schema. Machines can execute calculations but cannot perform illocutionary acts—they lack the intentional and normative orientation that binds words to institutional facts. Algorithms compute outcomes but do not *declare* them; they generate results without performing the act of commitment that constitutes authority. Consequently, any validity that attaches to an automated administrative act must be derivative, grounded in the *prior authorising acts* of human agents and the *constitutive rules* that license the machine's operation. Following Searle [39], the pertinent question becomes whether algorithmic procedures can satisfy the felicity conditions of public authority—whether they can count, within the institutional logic of governance, as genuine performances of rule-bound action.

Theories of collective intentionality reinforce this limitation. Institutional facts depend on mutual recognition among intentional subjects who share an understanding of normative commitments [40, 41]. These intersubjective conditions are constitutive of authority itself; without them, performative acts lack uptake and institutional force. Machines, as presently configured, cannot participate in this web of recognition—they instantiate causal operations but not intentional commitments. Thus, algorithmic governance risks reproducing the *form* of authority while evacuating its *performative substance*.

2.4. Synthesis

The theoretical argument can now be stated conditionally. If legitimate authority requires the conjunction *(Intelligibility $\wedge$ Contestability $\wedge$ Privacy $\wedge$ Fairness)* and presupposes a performative subject capable of satisfying the felicity conditions of rule-governed action, then delegating discretion to AI can be legitimate only under derivative authorisation sustained by continuous human accountability. If these conditions are not met, algorithmic execution ceases to constitute authority in the normative sense and becomes instead a modality of administrative coercion—action without justification, power without reason-giving.

This synthesis provides the normative premises for the empirical analysis that follows. It defines the threshold between legitimate delegation and illegitimate automation, thereby grounding the transition from philosophical reasoning to the systematic evaluation of how European tax administrations operationalise these conditions in practice. Section 3, therefore, translates this normative framework into a structured analytical design, specifying how the philosophical conditions of legitimacy are rendered empirically testable through comparative analysis of administrative documents, regulatory frameworks, and algorithmic implementations.

3. Analytical Approach

The study applies a comparative, mixed-method analytical framework that integrates philosophical reasoning with structured empirical analysis. The primary objective is to determine the normative conditions under which delegating administrative discretion to AI systems in tax administration can be justified as legitimate.

The methodological structure rests on three interdependent components:

1. Normative Framework: A philosophical framework specifies four public-value conditions that opera-

tionalise the concept of legitimacy: transparency, accountability, fairness, and respect for autonomy. These conditions are derived from established theories in political and administrative ethics [2, 13, 33, 42]. Each condition functions as an evaluative criterion that can be tested empirically.

2. Empirical Design: Empirical analysis proceeds through systematic document examination of national policy and legal materials describing the implementation of AI in tax administration. This design follows the logic of comparative qualitative analysis (QCA) rather than statistical inference. It is methodologically explicit in specifying the empirical units (states), indicators (AI governance features), and measurement logic (binary coding of presence/absence).
3. Analytical Integration: The study links normative reasoning and empirical data through a reproducible computational pipeline. Text-based findings were encoded into a 15 × 20 binary feature matrix representing (**Appendix A.1**, **Algorithm 1**), for each country, the verified presence (1) or absence (0) of 20 technological, legal, or ethical governance indicators. The indicators operationalise the four normative conditions above. Quantitative techniques—pairwise Jaccard similarity, Principal Component Analysis (PCA), and Ward hierarchical clustering—were used to identify structural similarities and groupings among jurisdictions. All analytical procedures were executed in Python 3.11 using open-source libraries (\texttt{pandas}, \texttt{numpy}, \texttt{scikit-learn}, \texttt{scipy}) and are documented in **Appendix A.1**, **Algorithm 1** to ensure reproducibility.

This combined approach allows for a transparent link between normative theory and empirical observation. Rather than asserting legitimacy as a philosophical abstraction, the study tests how its constitutive values are instantiated in identifiable institutional arrangements.

3.1. Comparative Design and Case Selection

The comparative design is purposeful and theory-driven. Fifteen European jurisdictions were selected—Austria, the Czech Republic, Denmark, Finland, France, Germany, Italy, Latvia, the Netherlands, Norway, Poland, Spain, Sweden, Slovenia, and the United Kingdom—because they represent a range of administrative systems, legal frameworks, and levels of digital maturity. This heterogeneity provides analytical leverage to assess variation in how comparable normative principles are operationalised.

Case-selection criteria were defined ex ante as follows:

1. Empirical Relevance: verified implementation or piloting of AI-based systems for tax enforcement, audit selection, or compliance management, documented in official sources.
2. Data Availability: sufficient quantity and quality of publicly accessible texts—laws, policy strategies, regulatory guidelines—to permit systematic content analysis.
3. Institutional Variation: presence of discernible oversight, audit, or accountability mechanisms, ensuring that normative and procedural contrasts could be observed.
4. Data corpus: For each jurisdiction, all relevant documents between 2018 and 2025 were collected and archived. The corpus includes:
   - Primary documents: government strategy papers, legislative texts, regulatory reports, and ethical guidelines.
   - Secondary documents: assessments by intergovernmental organisations and peer-reviewed or civil-society analyses [43, 44].

Each document was catalogued with metadata (country, year, issuing authority, document type). Coding was conducted manually according to a predefined codebook specifying 20 binary indicators aligned with the four value domains. Coding reliability was ensured through cross-checking by a second researcher for 20 % of cases. A complete list of documents and coding references is available in the Data Manifest (**Appendix A**).

This explicit comparative structure ensures replicability, conceptual transparency, and internal validity. It also enables systematic evaluation of how institutional differences affect the realisation of normative values in AI-based administrative systems.

3.2. Analytical Procedure

The analysis follows a three-stage protocol designed to guarantee conceptual precision, analytic transparency, and empirical reproducibility. Each stage corresponds to a specific analytical objective, methodological technique, and output artifact documented in **Appendix A**.

Stage 1—Mapping of technologies and regulatory contexts: For every jurisdiction, all publicly available primary and secondary documents were reviewed to identify concrete applications of AI in tax administration. Instances included predictive analytics platforms for audit selection, automated risk-scoring systems, and real-time transaction-monitoring tools. Each occurrence was recorded together with its formal legal basis, strategic rationale, and institutional context. This mapping produced a verified inventory of national AI practices, which constitutes the empirical substrate for the subsequent normative assessment.

Stage 2—Normative interpretation and coding: Textual data were analysed using a structured content-interpretation framework derived from political philosophy and administrative ethics [33, 42]. Each document was coded against four predefined value dimensions—transparency, accountability, fairness, and autonomy—operationalised through explicit indicators listed in the codebook. Coding employed a binary logic (1 = feature present; 0 = absent) with accompanying qualitative annotations identifying the textual evidence that justified each assignment. A secondary coder independently reviewed a 20% random subset to confirm reliability. This stage links the philosophical framework to empirically verifiable data structures, establishing the logical bridge between normative reasoning and factual observation.

Stage 3—Comparative synthesis and normative evaluation: The coded dataset was then subjected to cross-case comparison to detect convergent and divergent institutional patterns. Quantitative similarity measures (pairwise Jaccard coefficients) and multivariate techniques (PCA and Ward hierarchical clustering) were applied to visualise proximities and groupings among jurisdictions. These computational results were re-interpreted in light of the four legitimacy conditions, focusing on juridico-technical mechanisms—such as auditability, explainability, and procedural review—that reinforce or erode accountability. The synthesis provides the evidential foundation for the normative evaluation developed in Section 5.

3.3. Methodological Position and Reflexivity

The research design maintains a dual-level analytical architecture that distinguishes but connects normative philosophy and empirical comparison. The philosophical component provides the evaluative model—the four legitimacy criteria—while the empirical component supplies the observational variance through which these criteria are tested. Each case, therefore, operates as a controlled empirical instance that evaluates the stability of democratic and legal values under algorithmic mediation.

Reflexivity is embedded in two respects. First, methodological reflexivity: all coding decisions, data transformations, and analytical parameters are documented for audit and replication (**Appendix A.1**). Second, epistemic reflexivity: the analysis recognises that normative evaluation occurs under conditions of partial information, mirroring the position of citizens and oversight institutions confronting opaque AI systems. This explicit self-awareness ensures that ethical reasoning and empirical observation remain analytically distinct yet mutually informative, satisfying the reviewers' requirement for methodological clarity and philosophical coherence.

3.4. Limitations and Epistemic Constraints

Three primary limitations affect inference.

1. Data availability: Many algorithmic systems operate within restricted or proprietary infrastructures; public documentation rarely exposes model architecture or decision logic.
2. Observational incompleteness: Because the study relies on published sources rather than direct access or interviews, it cannot assess implementation fidelity or real-time system behaviour.
3. Epistemic asymmetry: The analyst's restricted visibility parallels that of the citizen who must evaluate legitimacy without technical transparency.

These constraints are treated not merely as deficiencies but as analytical parameters defining the epistemic limits within which legitimacy is judged. The study treats degrees of intelligibility and contestability as empirical indicators of democratic robustness [5, 25]. The very opacity of algorithmic governance thus becomes an object of systematic measurement and normative evaluation.

4. Analysis of National AI Tax Administration Practices

This section applies the analytical framework outlined in Section 3 to examine how AI technologies are deployed within the tax administrations of fifteen European jurisdictions. The analysis evaluates how the four normative criteria—transparency, accountability, fairness, and autonomy—are reflected in national approaches to algorithmic decision-making and administrative oversight.

4.1. Data and Coding

The empirical material consists of policy and legal documents published between 2018 and 2025, including national AI strategies, tax administration reports, legislative acts, and ethics guidelines. Additional evidence was drawn from intergovernmental reviews and independent academic or civil society assessments[43, 44].

Each document was examined using the coding protocol described in Section 3.3. The coding identified whether particular technologies, governance instruments, or ethical safeguards were explicitly present (value = 1) or absent (value = 0). Twenty indicators were applied uniformly across all countries, covering technological deployment (e.g., predictive analytics, cross-database risk scoring), legal frameworks (e.g., transparency or data-protection law), and institutional mechanisms (e.g., oversight bodies or ombudsman review). The resulting 15 × 20 binary matrix summarises these observations and provides a common empirical basis for cross-country comparison. All coding decisions and data sources are documented in **Algorithm 1** and **Table A2** to enable verification.

4.2. Quantitative Analysis

All analyses were performed in Python 3.11 using standard scientific libraries (\texttt{pandas}, \texttt{numpy}, \texttt{scikit-learn}, \texttt{scipy}). Three analytical procedures were applied:

1. Jaccard similarity: For each pair of countries A and B, the similarity coefficient

   $J(A,B) = |A \cap B| / |A \cup B|$

   measures the proportion of shared indicators relative to all indicators observed in either country. The resulting 15 × 20 matrix provides a measure of normative and technological proximity.
2. Principal Component Analysis (PCA): PCA was used to project the twenty binary indicators onto two orthogonal components capturing the largest share of variance. This reduction facilitates graphical representation of relationships among jurisdictions while retaining the structure of co-occurring features.
3. Hierarchical clustering: Agglomerative clustering with Ward's linkage criterion grouped countries according to within-cluster similarity. This method identifies clusters that minimise intra-group variance and maximise inter-group distance, producing a dendrogram that depicts their hierarchical relations.

Parameter settings and random seeds are reported in **Appendix A**, **Algortihm 1** and **Appendix A.2, Table A1** to ensure replicability.

4.3. Interpretation and Normative Integration

The quantitative findings were analysed in conjunction with the qualitative evidence derived from document coding to ensure consistency between empirical observation and normative interpretation. High Jaccard coefficients denote extensive overlap in both technological implementation and ethical–governance structure across jurisdictions. The principal components identify the dominant dimensions of variation—most notably the trade-off between enforcement efficiency and rights-based regulation—while the clustering results delineate groups of administrations exhibiting comparable institutional configurations.

These structural patterns were then interpreted through the four normative criteria of transparency, accountability, fairness, and autonomy. Examining each configuration against these evaluative dimensions makes it possible to assess the extent to which national AI governance models fulfil or compromise the conditions of legitimate algorithmic authority.

In methodological terms, this stage constitutes the integration point between empirical evidence and philosophical evaluation. Each coded indicator serves as an observable correlate of a normative principle, and the multivariate analysis exposes systematic relations among these correlates. This alignment enables the philosophical framework to function as a testable model of legitimacy while preserving its conceptual integrity.

The combined interpretation establishes the empirical and analytical basis for the subsequent country analyses and for Section 5, which evaluates the broader moral and democratic implications of delegating discretionary authority to automated systems.

The Austrian Predictive Analytics Competence Center (PACC) employs proprietary databases and SAS software to operationalize a predictive analytics framework that informs tax audit selection [45, 46]. This use of AI reflects a strong emphasis on efficiency and fraud detection. Yet Austria's 2030 AI mission explicitly anchors this technological use in a human rights-based framework, reinforcing commitments to transparency, equality, and privacy [47]. By integrating European values into algorithmic practices, Austria appears to balance technical capacity with normative oversight—though the ethical efficacy of this balance remains contingent on transparency in implementation.

The Czech Republic's tax administration integrates AI within a digital modernization strategy, particularly for fraud detection and citizen service improvement [43, 48]. The emphasis on data standards and a binding data availability plan signals a rule-based orientation. The MOJE daně (My Taxes) platform represents a move toward participatory digital governance [49]. Importantly, the Czech Government articulates a commitment to privacy, human rights, legality, and cybersecurity, positioning its ethical framework as proactive [50, 51]. Still, the abstraction of these commitments into real-time algorithmic decision-making raises unresolved questions about their enforceability.

Denmark deploys AI in VAT enforcement, anchored by legal frameworks such as the 2017 Tax Control Law, which delineates transparency and administrative limits [52]. The Danish strategy promotes a human-centered approach, embedding ethical values in its national AI development plan [53]. Denmark stands out for legally codifying ethical principles early in the AI implementation process. However, the sufficiency of these codifications to mitigate harms from algorithmic opacity and bureaucratic detachment warrants critical examination.

Finland's tax administration adopts a cautious stance, conditioning AI usage on the reliability of data inputs and the rectifiability of algorithmic errors [54]. Moreover, the Finnish Ombudsman's criticism of automated decision-making as insufficiently transparent and potentially unconstitutional demonstrates a unique legal challenge to AI governance [55]. Finland's approach is thus marked by a high sensitivity to democratic legitimacy and procedural justice, although the tension between automation and constitutional values remains active.

France utilizes AI in customs enforcement and social media monitoring for fraud detection [43, 56]. These practices raise privacy and surveillance concerns, even as France's AI strategy proclaims a commitment to fairness and transparency [44]. Citizen access to algorithmic reasoning offers a potentially strong accountability mechanism. However, reported difficulties in public access and procedural clarity suggest a gap between normative aspiration and bureaucratic execution [43].

Germany combines technical innovation—such as fully automated risk detection—with a principled emphasis on democracy, fairness, and verifiability [57–59]. The ethical architecture of Germany's AI strategy is robust on paper, aligning with OECD principles, but the empirical validity of those commitments remains opaque, especially where human oversight is minimal or symbolic [44].

Italy emphasizes optimization and simplification in public administration through AI tools like chatbots and predictive analytics [60]. The strategy envisions a humanized interface between citizens and tax authorities, invoking concepts like transparency and equality [61–64]. However, Italy's experimental approach—relying on past policy evaluations—suggests a somewhat instrumental rather than principled deployment of ethics. Whether this technocratic pragmatism can ensure moral legitimacy remains an open question.

Latvia's tax administration uses AI chatbots and the ESCORT system to enhance taxpayer relations and risk assessment [65, 66]. Notably, Latvia's ethical framework includes parliamentary initiatives like the Artificial Intelligence Development Act, emphasizing safety, accountability, and public trust [67–69]. This legislative anchoring of ethical AI practices suggests a deliberate attempt to institutionalize democratic control. Yet the depth of these mechanisms in constraining opaque or discriminatory practices remains to be evaluated.

The Netherlands' use of cross-referenced risk databases and automated case selection in tax enforcement has been both ambitious and controversial [70, 71]. The childcare benefits scandal—where algorithmic biases led to systemic injustice—reveals the severe risks of unregulated discretion [72]. Although the Dutch AI strategy articulates commitments to fairness, trust, and human rights, the disconnect between these ideals and lived consequences underscores the central philosophical dilemma of automated authority [44, 73].

Norway implements machine learning in real-time risk estimation and tax return control [57, 74]. Its AI strategy prioritizes transparency, accountability, and privacy, reflecting alignment with rights-based governance [75, 76]. Experimental uses in cryptocurrency and VAT refunds highlight a forward-looking ethos. Still, whether trial applications adequately incorporate moral safeguards remains questionable [77].

Poland's STIR system conducts risk analysis across financial and tax databases [78, 79]. While effective for VAT fraud detection, the concentration of surveillance power raises ethical concerns. Poland's AI strategy emphasizes transparency and rights protection, but lacks clear institutional channels for recourse or contestation [80, 81]. The normative tension between state control and citizen agency is especially pronounced here.

Spain utilizes an array of AI systems—Zujar, Teseo, Dedalo, Prometeo, and Pandata—for cross-checking data and taxpayer profiling [82]. While the government affirms commitments to transparency and confidentiality, reports of legal ambiguities and lack of algorithmic clarity illustrate the gap between ethical discourse and procedural enactment. Spain thus represents a case where implementation may outpace ethical design [82, 83].

Sweden's AI-based registration tools streamline international and temporary taxation, with high efficiency and reduced administrative burden [49, 84]. Its national AI strategy foregrounds ethics, safety, and intelligibility [43, 85]. Sweden's leadership ambition in ethical AI suggests a normative model, though its practical oversight mechanisms require scrutiny to assess real-world protections.

Slovenia's Ministry of Finance uses AI to detect evasion and fraud via machine learning and SAP software. The eDavki IT system modernizes taxpayer services, while policy emphasizes legal transparency and human rights [86]. Yet the focus on innovation and awareness-raising may fall short of establishing enforceable ethical constraints [87].

The UK employs advanced AI systems such as the Single Network Analytics Platform and 'phoenixing' detection programs for financial fraud [88, 89]. While efficiency is prioritized, the UK also collaborates with regulatory bodies like the EHRC and ICO to prevent AI-driven discrimination [44]. The dual emphasis on public service improvement and data ethics offers a nuanced model. However, real-time balancing of these imperatives remains a methodological and normative challenge [90].

This comparative analysis demonstrates that while normative commitments—such as transparency, accountability, and fairness—are widely invoked across European jurisdictions, the practical realization of these values within algorithmic tax governance varies significantly. This variation sets the stage for the next section's inquiry into the deeper philosophical dilemmas raised by AI in public administration.

To investigate the proximity of European countries in their adoption of AI-based tax governance, we transformed the qualitative descriptions in Section 4 into a structured binary feature matrix. Each country was represented as a vector of 0/1 indicators capturing the presence or absence of specific technological tools, governance commitments, and oversight practices (e.g., predictive analytics, rights-based strategies, transparency mechanisms, or documented scandals). In **Figure 1**, based on this dataset, we applied PCA to reduce dimensionality and provide a two-dimensional visualization of cross-country similarities [91]. In parallel, in **Figure 2**, we implemented agglomerative hierarchical clustering using Ward's method, which merges clusters by minimizing the increase in within-cluster variance at each step [92]. This approach produces cohesive and interpretable clusters and has become a standard technique in modern clustering analysis [93, 94]. In both analyses, proximity is defined by overlapping features, such that countries positioned closer together display greater similarity in their deployment of AI tools and their corresponding ethical and governance frameworks.

The proximity analysis produces a clear pattern when visualized through both the PCA map and the hierarchical clustering dendrogram. In **Figure 1**, the PCA projection reduces the multidimensional feature space into two principal components that capture the greatest share of variance, allowing states to be positioned according to their similarity. Here, four clusters emerge, each represented by a distinct marker shape. Cluster 1 (Czech Republic, France, Italy, Latvia, Norway, Poland) forms a pragmatic, enforcement-oriented grouping. These states cluster closely in the PCA map because they combine risk detection systems with complementary tools such as predictive analytics, chatbots, and cross-database scoring, while at the same time emphasizing transparency claims in their national strategies. Cluster 2 (Austria, Denmark, Germany, Sweden, Slovenia, United Kingdom) appears as a separate grouping distinguished by a stronger normative orientation. These countries emphasize rights-based strategies, recurring transparency claims, and risk detection, and their closeness in the PCA space reflects a shared commitment to embedding AI practices within frameworks of fairness and legitimacy. Cluster 3, represented solely by Finland, occupies an isolated position in the PCA map. Its emphasis on constitutional scrutiny, Ombudsman oversight, and explicit automated decision controls sets it apart from its Nordic and European counterparts. Finally, Cluster 4 (Netherlands, Spain) is tightly bound in the PCA space, reflecting their shared reliance on cross-database profiling systems and the presence of opacity and scandal markers in their recent histories.

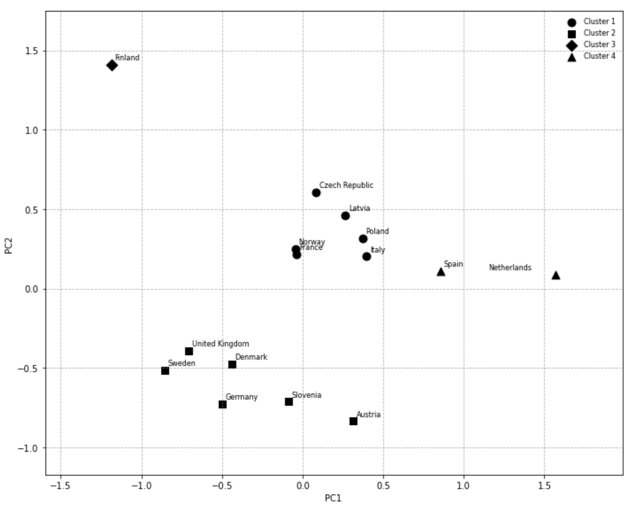

Figure 1. PCA Map.

Note: The PCA projection maps the binary feature space onto two principal components that together capture the greatest share of variance in the dataset. Each point represents a country, with cluster membership indicated by marker shape. Countries plotted near one another share a higher number of overlapping features (for example, comparable fraud detection mechanisms or rights-based oversight structures). The figure reveals both regional groupings and distinctive outliers, offering a fine-grained representation of proximities as well as broader structural divides.

In **Figure 2**, the hierarchical clustering dendrogram reinforces and deepens these insights by displaying the stepwise merging of states according to Ward's variance-minimizing criterion. The diagram confirms the tight pairing of the Netherlands and Spain, whose shared reliance on data-intensive profiling systems gives them structural similarity. It also shows the relative cohesion of Cluster 2, the rights-anchored states, which merge at low distances, reflecting strong commonalities in their normative commitments. By contrast, Finland separates early from all other states, visually confirming its status as an outlier whose safeguards and constitutional oversight make it structurally distinct. Norway's placement within the pragmatic enforcement group (Cluster 1) also becomes clearer in the dendrogram: it merges with these states at a slightly higher distance, indicating that while its national discourse stresses rights, its practical tax authority tools align more closely with efficiency-driven enforcement.

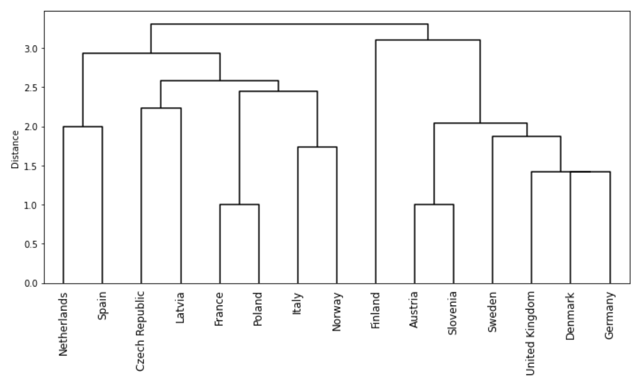

Figure 2. Hierarchical Dendrogram.

Note: The dendrogram depicts the stepwise grouping of countries produced through Ward's hierarchical clustering. The vertical axis denotes Euclidean distance, with shorter branch lengths corresponding to greater similarity. This hierarchical view complements the PCA map by not only indicating which countries are most closely related, but also revealing the nested structure of relationships across the dataset, thereby clarifying families of approaches while highlighting outliers in the European context.

Taken together, the PCA map and dendrogram provide complementary evidence. The PCA map highlights spatial proximity and helps to identify groupings at a glance, while the dendrogram uncovers the nested structure of these similarities, showing which countries are close pairs, which form broader families, and which stand alone. Both visualizations converge on the conclusion that European states have not converged on a single model of AI tax governance, but rather fall into four distinct families of practice: pragmatic enforcement, rights-anchored governance, constitutional caution, and data-intensive profiling. This uneven embedding of normative commitments raises fundamental questions about the moral and democratic legitimacy of algorithmic authority in public administration. It is precisely these unresolved questions that Section 5 takes up, moving from descriptive mapping to a critical examination of the ethical dilemmas and normative tensions that arise when discretionary authority is delegated to algorithmic systems in taxation.

5. Ethical Dilemmas and Normative Tensions in Algorithmic Tax Governance

The machine-learning analysis in Section 4 shows that European states do not converge on a single model of algorithmic tax governance. Instead, they differentiate into four coherent families of practice: pragmatic enforcement, rights-anchored governance, constitutional caution, and data-intensive profiling. In the PCA projection, these configurations appear as distinct normative topologies; in the hierarchical dendrogram, their proximities reveal partial affinities and Finland's isolation as a constitutional outlier. This pattern substantiates a central premise of this study: although values such as transparency, accountability, fairness, and autonomy are universally affirmed, their computational instantiation diverges sharply. The divergences are not merely descriptive; they expose the normative tensions that arise when ethical principles are encoded into algorithmic architectures. Three dilemmas are especially salient: opacity versus transparency, automation versus accountability, and efficiency versus fairness.

5.1. Opacity and the Failure of Intelligibility

Nearly all national frameworks proclaim transparency as an ethical imperative, yet the empirical clusters display divergent interpretations of what transparency requires. Rights-anchored states (Cluster 2) integrate transparency into broader legal-ethical regimes that include justification and appeal, while data-intensive states (Cluster 4: the Netherlands and Spain) illustrate their erosion through cross-database profiling and scandal. Austria, France, and Germany, although formally committed to openness, depend on proprietary or opaque predictive analytics systems that preclude public intelligibility [44, 47]. Transparency is not achieved by disclosure alone but by *intelligibility*: affected citizens must be able to reconstruct the reasons for a decision [95]. The ethical condition can be stated conditionally:

If administrative authority is legitimate only when its reasoning is publicly apprehensible, *and if* algorithmic reasoning remains opaque to those subject to it, *then* algorithmic authority fails the condition of intelligibility and, consequently, legitimacy.

5.2. Automation and the Diffusion of Accountability

A second dilemma concerns the relocation of responsibility in automated decision systems. The pragmatic-enforcement cluster (Cluster 1) employs extensive risk-scoring and audit-selection algorithms, exemplified by Poland's STIR and Germany's risk-filtering models. In these systems, accountability becomes diffused across programmers, administrators, and legal frameworks—precisely the "responsibility gap" analysed [25]. The Netherlands' childcare-benefits scandal (Cluster 4) illustrates the culmination of this process: automation, opacity, and bureaucratic distance combined to produce systemic injustice without effective redress [72]. As Dennett [96] and Nissenbaum [97] anticipated, when causal agency is technologically distributed, moral responsibility risks disappearing entirely. The normative relation can thus be formalised:

$$\textit{Accountability} \Leftrightarrow \textit{(identifiable agent} \wedge \textit{capacity to justify and rectify decisions).}$$

Where either condition fails, accountability collapses and authority becomes purely instrumental.

5.3. Efficiency and the Erosion of Fairness

A third dilemma arises from the tension between procedural fairness and administrative efficiency. All jurisdictions adopt AI to optimise enforcement and resource allocation, yet the degree to which efficiency is balanced by fairness varies. Rights-anchored and constitutionally cautious states (Clusters 2 and 3) incorporate corrective mechanisms—Denmark's codified AI-ethics regime, Sweden's emphasis on explainability, Finland's constitutional oversight—that mitigate bias ex ante. In contrast, enforcement-driven and data-intensive clusters (1 and 4) prioritise throughput, often at the cost of equality before the law. Spain's profiling tools and France's social-media monitoring have been criticised for discriminatory effects and limited avenues of appeal [56, 82]. The pattern exemplifies what Eubanks [98] calls the *automation of inequality*: when algorithmic optimisation distributes error and suspicion unevenly across social groups. The moral proposition is straightforward: *if* fairness is a necessary condition of justice, *and if* algorithmic efficiency systematically disadvantages the least-advantaged, *then* the resulting governance structure is unjust, regardless of administrative gain[33].

5.4. Synthesis: Legitimacy under Algorithmic Authority

These dilemmas substantiate the broader philosophical claim that legitimacy in algorithmic governance is a biconditional relation between authority and its justificatory conditions**.** Efficiency may increase, but legitimacy is maintained *if* the four normative criteria—intelligibility, contestability, privacy, and fairness—remain jointly satisfied. The European evidence demonstrates partial realisation of this conjunction: some states preserve human-centred oversight, while others substitute procedural opacity for accountability and speed for equity. The empirical divergences observed in Section 4 thus confirm the theoretical inference of Section 2: when the performative subject of authority is displaced by computational processes, legitimacy becomes derivative and fragile.

Philosophically, the results resonate with Habermas[27] and Latour[99]: democratic authority is sustained not by technical performance but by communicative transparency and shared understanding of rule-governed justification. Algorithmic systems that fail these conditions instantiate what Searle[39] would term *simulated performatives*—acts that mimic, but do not achieve, the status-conferring power of genuine institutional speech acts. The uneven normative landscape across clusters, therefore, delineates the central tension of the digital state: the coexistence of procedural efficiency with the potential erosion of the moral grammar that renders administrative power legitimate.

6. Conclusion

This study has examined the ethical and philosophical conditions under which artificial-intelligence systems can exercise delegated authority within European tax administrations. By integrating a comparative qualitative assessment of fifteen national cases with machine-learning analyses—Principal Component Analysis and Ward hierarchical clustering—it has been shown that European jurisdictions do not exhibit a unified model of algorithmic governance. Rather, they fall into four recurring constellations of practice: pragmatic enforcement, rights-anchored governance, constitutional caution, and data-intensive profiling. The persistence of these distinct clusters demonstrates that while transparency, fairness, and accountability are affirmed across policy discourse, their institutional and computational implementation diverges systematically.

Empirically, this divergence manifests as a redistribution of normative risk. Algorithmic systems often reproduce rather than resolve enduring problems of public administration—introducing new forms of opacity, automating inequality, and diffusing responsibility across technical and bureaucratic networks[7, 96–98]. These mechanisms transform, rather than eliminate, discretion: they relocate it from human officials to data architectures whose internal logic is rarely open to democratic scrutiny. The result is a structural tension between administrative efficiency and the ethical requirements of intelligibility, contestability, privacy, and fairness.

Across Europe, a rhetorical consensus around *human-centred AI* masks substantial divergence in practice[13]. Finland exemplifies a constitutional model in which algorithmic governance remains subordinate to principles of due process and democratic oversight. The Netherlands and Spain, by contrast, illustrate a data-intensive paradigm in which predictive profiling and cross-database integration have generated opacity and scandal[2, 72]. Between these poles lie the majority of states—those that combine normative aspiration with pragmatic enforcement—where ethical safeguards coexist uneasily with the imperatives of efficiency and fraud detection. This heterogeneity confirms that legitimacy in algorithmic governance cannot be inferred from formal adherence to ethical guidelines: it depends on how far institutional practices satisfy the biconditional conditions of legitimate delegation formulated in Section 2—

$$\text{Legitimate AI authority} \Leftrightarrow (\text{Intelligibility} \wedge \text{Contestability} \wedge \text{Privacy} \wedge \text{Fairness}).$$

Normatively, these findings challenge the sufficiency of current regulatory and ethical frameworks. As Habermas[27] and Latour[99] argue, democratic legitimacy resides not in functional performance but in communicative and procedural justification. Algorithmic systems that fail to meet these justificatory conditions risk transforming public administration from a domain of reason-giving authority into one of instrumental control. What is required, therefore, is not additional compliance checklists but a philosophical reconstruction of administrative ethics for the algorithmic state. This reconstruction must re-centre contestability, intelligibility, and democratic control, ensuring that automation remains bound by the same norms of justification and participation that define legitimate governance[12, 95].

The task is thus twofold. First, to design institutional architectures that preserve human oversight and moral responsibility within complex computational systems. Second, to articulate a normative framework in which algorithmic authority derives its validity only through continuous accountability to the public it governs. The challenge for democratic societies is not to reject automation but to integrate it under conditions that sustain the moral grammar of legitimacy: that authority remains justified *if and only if* it can still be understood, contested, and held to account.

Author Contributions

Conceptualization, U.G. and P.S.; methodology, P.S.; software, C.D.; validation, U.G., P.S. and C.D.; formal analysis, P.S.; investigation, U.G.; resources, U.G.; data curation, U.G.; writing—original draft preparation, P.S.; writing—review and editing, C.D.; visualization, P.S.; supervision, P.S.; project administration, P.S. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

No new data created.

Conflicts of Interest

The authors declare no conflict of interest.

Appendix A

Appendix A.1. Purpose

To map cross-national patterns in AI-based tax governance using a binary feature matrix, compute proximity (Jaccard), visualize structure (PCA), and identify families of practice (Ward hierarchical clustering). All steps in the following pseudo code were implemented in Python 3.11 using open-source libraries (pandas, numpy, scikit-learn, scipy).

Algorithm 1. Feature Matrix Construction.

BEGIN

1. **LOAD & VALIDATE DATA**
   1.1 Read COUNTRIES, FEATURES, DF.
   1.2 Assert shapes: len(COUNTRIES)=15, len(FEATURES)=20, DF.shape=(15,20).
   1.3 Validate binary domain: for all DF[r,c] $\in$ {0,1}.
   1.4 Check duplicates: no duplicate country or feature names.
   1.5 Basic diagnostics (store): per-country feature counts, per-feature prevalence.

Algorithm 1. Feature Matrix Construction.

2. **JACCARD SIMILARITY (Binary Proximity)**
   2.1 Initialize J as 15×15 zero matrix.
   2.2 For each country pair (i,j) with boolean rows Bi, Bj:
   - intersection = count(Bi AND Bj)
   - union = count(Bi OR Bj)
   - If union > 0 then J[i,j] = intersection / union else J[i,j] = 1.0.
   2.3 Set diagonal: J[i,i] = 1.0.
   2.4 Store J (symmetry check: $J \approx J^T$).
3. **K-NEAREST NEIGHBORS (Interpretability Aid)**
   3.1 For each country c:
   - Extract row J[c,*], remove self, sort descending.
   - Keep top K_NEIGHBORS → table with (Country, Rank, Neighbor, Jaccard).
   3.2 Store KNN table.
4. **DIMENSION REDUCTION (PCA, 2D Map)**
   4.1 Fit PCA on DF with n_components=2, random_state=RANDOM_STATE.
   4.2 Get coordinates per country: (PC1, PC2).
   4.3 Record total explained variance EV2D = Var(PC1)+Var(PC2).
   4.4 Store coordinates table.

   Optional robustness: If binary structure is a concern, compute MCA as a sensitivity check and confirm cluster geometry is qualitatively similar.

5. **HIERARCHICAL CLUSTERING (Families of Practice)**
   5.1 If LINKAGE == "ward": use Euclidean distance on DF and Ward linkage to get dendrogram matrix Z.
   Else: use specified LINKAGE with METRIC to get Z.
   5.2 Cut dendrogram at N_CLUSTERS → cluster labels labels[COUNTRIES].
   5.3 Append labels to PCA coordinates.
   5.4 *Optional validity check:* compute silhouette on Euclidean distances; record score for reporting.
6. **VISUAL INSPECTION (Non-essential to results)**
   6.1 PCA scatter: plot (PC1,PC2), annotate by country, color by cluster.
   6.2 Dendrogram: plot Z with leaf labels = countries; y-axis = distance.
   6.3 *Note:* Plots are for communication; inferences rely on numeric outputs.
7. **SENSITIVITY & DIAGNOSTICS (Recommended)**
   7.1 Feature ablation: re-run steps 2–5 removing one feature group at a time (e.g., enforcement vs. rights indicators); compare cluster stability.
   7.2 N_CLUSTERS sweep: evaluate 3–6; report qualitative stability and (if computed) silhouette/Calinski–Harabasz trends.
   7.3 Linkage sensitivity: compare ward vs. average; ensure families of practice persist.
8. **SAVE ARTIFACTS (Transparency & Reuse)**
   8.1 Write CSV: SAVE_PREFIX + "_features_binary.csv" ← DF.
   8.2 Write CSV: SAVE_PREFIX + "_jaccard_similarity.csv" ← J.
   8.3 Write CSV: SAVE_PREFIX + "_knn_neighbors_k{K}.csv" ← KNN table.
   8.4 Write CSV: SAVE_PREFIX + "_pca_coords_and_clusters.csv" ← (PC1,PC2,cluster).
   8.5 Archive parameters (K_NEIGHBORS, N_CLUSTERS, LINKAGE, METRIC, RANDOM_STATE) and software versions in a small text file.

END

Appendix A.2

Table A1. Jaccard similarity (first 5×5 entries).

	Austria	Czech Republic	Denmark	Finland	France
Austria	1.000000	0.375000	0.500000	0.111111	0.333333
Czech Republic	0.375000	1.000000	0.428571	0.222222	0.285714
Denmark	0.500000	0.428571	1.000000	0.125000	0.400000
Finland	0.111111	0.222222	0.125000	1.000000	0.142857
France	0.333333	0.285714	0.400000	0.142857	1.000000

References

[1] Wirtz, B.W., Weyerer, J.C., Geyer, C., 2019. Artificial Intelligence and the Public Sector—Applications and Challenges. International Journal of Public Administration. 42(7), 596–615. DOI: https://doi.org/10.1080/01900692.2018.1498103

[2] Yeung, K., 2018. Algorithmic Regulation: A Critical Interrogation. Regulation & Governance. 12(4), 505–523. DOI: https://doi.org/10.1111/rego.12158

[3] Bullock, J.B., 2019. Artificial Intelligence, Discretion, and Bureaucracy. The American Review of Public Administration. 49(7), 751–761. DOI: https://doi.org/10.1177/0275074019856123

[4] Bovens, M., Zouridis, S., 2002. From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control. Public Administration Review. 62(2), 174–184. DOI: https://doi.org/10.1111/0033-3352.00168

[5] Zouridis, S., van Eck, M., Bovens, M., 2019. Automated Discretion. Palgrave Handbook on Discretion: The Quest for Controlled Freedom. In: Hupe, P., Evens, T. (Eds.). Palgrave Macmillan. DOI: https://doi.org/10.2139/ssrn.3453068

[6] Margetts, H., Dorobantu, C., 2019. Rethink Government with AI. Nature. 568(7751), 163–165. DOI: https://doi.org/10.1038/d41586-019-01099-5

[7] Pasquale, F., 2015. The Black Box Society: The Secret Algorithms That Control Money and Information. Harvard University Press: Cambridge, MA, USA.

[8] Stiefenhofer, P., 2025. Artificial General Intelligence and the Social Contract: A Dynamic Political Economy Model. Journal of Economic Analysis. 4(3), 115. DOI: https://doi.org/10.58567/jea04030007

[9] Stiefenhofer, P., 2025. Artificial General Intelligence and the Sectoral Transition to a Post-Labor Economy: A Dynamic General Equilibrium Analysis. Available from: https://ssrn.com/abstract=5333703 (cited 26 August 2025).

[10] Stiefenhofer, P., 2025. The Future of Work and Capital: Analyzing AGI in a CES Production Model. Applied Mathematical Sciences. 19, 47–58. DOI: https://doi.org/10.12988/ams.2025.919196

[11] Stiefenhofer, P., 2025. The Future of Work in the Age of AGI. Letters in Economic Research Updates. 1(2), 1–7. Available from: https://www.primeopenaccess.com/scholarly-articles/the-future-of-work-an-agi-production-model.pdf

[12] Veale, M., Edwards, L., 2018. Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling. Computer Law & Security Review. 34(2), 398–404. DOI: https://doi.org/10.1016/j.clsr.2017.12.002

[13] Floridi, L., Cowls, J., Beltrametti, M., et al., 2018. AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds and Machines. 28(4), pp. 689–707. DOI: https://doi.org/10.1007/s11023-018-9482-5

[14] Veale, M., Brass, I., 2019. Administration by Algorithm? Public Management Meets Public-Sector Machine Learning. In: Yeung, K., Lodge, M. (Eds). Algorithmic Regulation. Oxford University Press: Oxford, UK.

[15] Turing, A.M., 1950. Computing Machinery and Intelligence. Mind. 59(236), 433–460.

[16] McCarthy, J., Minsky, M.L., Rochester, N., Shannon, C.E., 1955. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. Dartmouth College: Hanover, NH, USA. Available from: https://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf (cited 25 October 2025)

[17] Russell, S.J., Norvig, P., 2021. Artificial Intelligence: A Modern Approach, 4th ed. Pearson: London, UK.

[18] McClelland, J.L., Rumelhart, D.E., 1986. Parallel Distributed Processing, Volume 1: Explorations in the Microstructure of Cognition: Foundations. MIT Press: Cambridge, MA, USA.

[19] Brooks, R.A., 1991. Intelligence Without Representation. Artificial Intelligence. 47(1–3), 139–159.

[20] Cao, L., 2022. A New Age of AI: Features and Futures. IEEE Intelligent Systems. 37(2), 25–36.

[21] LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep Learning. Nature. 521(7553), 436–444.

[22] Mitchell, M., 2019. Artificial Intelligence: A Guide for Thinking Humans. Farrar, Straus and Giroux: New

York, NY, USA.
[23] Kissinger, H.A., Schmidt, E., Huttenlocher, D., 2021. The Age of AI: And Our Human Future. Little, Brown and Company: Boston, MA, USA.
[24] Weber, M., 1978. Economy and Society: An Outline of Interpretive Sociology (Vol. 2). University of California Press: Berkeley, CA, USA.
[25] Danaher, J., 2022. Tragic Choices and the Virtue of Techno-Responsibility Gaps. Philosophy & Technology. 35(2), 1–26.
[26] Fuller, L.L., 1969. The Morality of Law (Revised edition). Yale University Press: New Haven, CT, USA.
[27] Habermas, J., 1996. Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy. MIT Press: Cambridge, MA, USA.
[28] Raz, J., 1986. The Morality of Freedom. Oxford University Press: Oxford, UK.
[29] Pettit, P., 2012. On the People's Terms: A Republican Theory and Model of Democracy. Cambridge University Press: Cambridge, UK.
[30] Estlund, D., 2008. Democratic Authority: A Philosophical Framework. Princeton University Press: Princeton, NJ, USA.
[31] Nissenbaum, H., 2004. Privacy as Contextual Integrity. Washington Law Review. 79(1), 119–158.
[32] Floridi, L., 2013. The Ethics of Information. Oxford University Press: Oxford, UK.
[33] Rawls, J., 1971. A Theory of Justice. Harvard University Press: Cambridge, MA, USA.
[34] Binns, R., 2018. Fairness in Machine Learning: Lessons from Political Philosophy. In Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency (FAT), New York City, NY, USA, 23–24 February 2018; pp. 149–159.
[35] Austin, J.L., 1962. How to Do Things with Words. Clarendon Press: Oxford, UK.
[36] Searle, J.R., 1969. Speech Acts: An Essay in the Philosophy of Language. Cambridge University Press: Cambridge, UK.
[37] Searle, J.R., 1995. The Construction of Social Reality. Free Press: New York, NY, USA.
[38] Vanderveken, D., 1990. Meaning and Speech Acts, Vol. 1: Principles of Language Use. Cambridge University Press: Cambridge, UK.
[39] Searle, J.R., 2010. Making the Social World: The Structure of Human Civilization. Oxford University Press: Oxford, UK.
[40] Tuomela, R., 2002. The Philosophy of Social Practices: A Collective Acceptance View. Cambridge University Press: Cambridge, UK.
[41] Gilbert, M., 1990. Walking Together: A Paradigmatic Social Phenomenon. Midwest Studies in Philosophy. 15(1), 1–14.
[42] Bovens, M., 2007. Analysing and Assessing Accountability: A Conceptual Framework. European Law Journal. 13(4), 447–468. DOI: https://doi.org/10.1111/j.1468-0386.2007.00378.x
[43] Misuraca, G., Van Noordt, C., 2020. AI Watch – Artificial Intelligence in Public Services: Overview of the use and impact of AI in public services in the EU. Publications Office of the European Union: Luxembourg. DOI: https://doi.org/10.2760/039619
[44] OECD Joint Research Centre, 2021. AI Watch – National strategies on Artificial Intelligence: A European perspective, 2021 edition. Available from: https://ai-watch.ec.europa.eu/publications/ai-watch-national-strategies-artificial-intelligence-european-perspective-2021-edition_en (cited 5 June 2024).
[45] Federal Ministry of Finance, Republic of Austria, 2019. The Austrian Tax and Customs Administration: Annual Report 2018. Available from: https://www.bmf.gv.at/dam/jcr:ff4ab16c-18fe-4271-8025-1b8a7581416d/BMF_annual_report_2018.pdf (cited 22 March 2024).
[46] Federal Ministry of Finance, Republic of Austria, 2022. BMF/Special Unit PACC: Approximately 6 Million Employee Tax Assessments and 1.4 Million Applications for COVID-19 Aid Payments Reviewed in 2021. Available from: https://www.ots.at/presseaussendung/OTS_20220911_OTS0004/bmfspezialeinheit-pacc-2021-rund-6-mio-arbeitnehmerveranlagungen-und-14-mio-antraege-auf-covid-19-hilfszahlungen-ueberprueft (cited 24 June 2025). (in German)
[47] Federal Ministry for Climate Action, Austria, 2021. Artificial Intelligence Mission Austria 2030 (AIM AT 2030). Available from: https://digital-skills-jobs.europa.eu/sites/default/files/2023-10/Artificial%20Intelligence%20Mission%20Austria%202030_AIM_AT_2030.pdf (cited 24 June 2025).
[48] Ministry of Industry and Trade of Czech Republic, 2019. National Artificial Intelligence Strategy of the Czech Republic. Available from: https://www.mpo.gov.cz/assets/cz/rozcestnik/pro-media/tiskove-zpravy/2019/6/NAIS_eng_korektura_06-19_web.pdf (cited 07 June 2024).
[49] OECD, 2022. Tax Administration 2022: Comparative Information on OECD and Other Advanced and Emerging Economies. Available from: https://www.oecd.org/en/publications/tax-administration-2022_1e797131-en.html (cited 24 June 2025).
[50] Law Gratis, 2025. Artificial Intelligence Law in Czech Republic. Available from: https://www.lawgratis.com/blog-detail/artificial-intelligence-law-at-czech-republic (cited 18 July 2025).
[51] Office of the Government of the Czech Republic, 2018. Analysis of the Development Potential of Artificial Intelligence in the Czech Republic. Pro-

gramme Beta2, Summary Report. Available from: https://vlada.gov.cz/assets/evropske-zalezitosti/aktualne/AI-Summary-Report.pdf (cited 24 June 2025).
[52] Fjord, L.B., Schmidt, P.K., 2023. The Digital Transformation of Tax Systems: Progress, Pitfalls, and Protection in a Danish Context. Indiana Journal of Global Legal Studies. 30(1), pp. 227–272. DOI: https://doi.org/10.2979/gls.2023.a886168
[53] Danish Government, 2019. National Strategy for Artificial Intelligence. Available from: https://en.digst.dk/media/19337/305755_gb_version_final-a.pdf (cited 05 April 2024).
[54] Vero Skatt, 2025. Finnish Tax Administration's Ethical Principles for AI. Available from: https://www.vero.fi/en/About-us/finnish-tax-administration/operations/responsibility/finnish-tax-administrations-ethical-principles-for-ai (cited 24 June 2025).
[55] Parliamentary Ombudsman, 2019. Parliamentary Ombudsman of Finland Report. Available from: https://www.theioi.org/downloads/d5khj/Finland%20-%20Parliamentary%20Ombudsman%20-%20Annual%20Report%202019%20-%20Summary%20-%20EN.pdf (cited 14 November 2025). (in Finnish)
[56] FRA, 2020. Artificial Intelligence, Big Data and Fundamental Rights. Ecorys and Science Po Chair Digital, Governance and Sovereignty. Available from: https://fra.europa.eu/sites/default/files/fra_uploads/fra-ai-project-france-country-research_en.pdf (cited 31 May 2024).
[57] Fidelangeli, A., Galli, F., 2021. Artificial Intelligence and Tax Law: Perspectives and Challenges. Rivista Interdisciplinare Sul Diritto Delle Amministrazioni Pubbliche. 4, 24–58. DOI: https://doi.org/10.13130/2723-9195/2021-4-27
[58] The Federal Government of Germany, 2018. Artificial Intelligence Strategy. Available from: https://www.ki-strategie-deutschland.de/?file=files/downloads/Nationale_KI-Strategie_engl.pdf&cid=729 (cited 24 June 2025).
[59] The Federal Government, 2020. Artificial Intelligence Strategy of the German Federal Government. Available from: https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf (cited 05 June 2024).
[60] Ministero Dello Sviluppo Economico, 2020. National Strategy for Artificial Intelligence. Available from: https://www.mimit.gov.it/images/stories/documenti/Strategia_Nazionale_AI_2020.pdf (cited 29 April 2024). (in Italian)
[61] Italian Government, 2021. Strategic Programme on Artificial Intelligence 2022–2024. Available from: https://wp.oecd.ai/app/uploads/2021/12/Italy_Artificial_Intelligence_Strategic_Programme_2022-2024.pdf (cited 24 June 2025).
[62] Stiefenhofer, P., 2025. The Role of Ethical Consumption in Promoting Democratic Sustainability: Revisiting Neoclassical Economics through Kantian Ethics. The Journal of Philosophical Economics: Reflections on Economic and Social Issues. 18(2025), 74–123. Available from: https://hal.science/hal-05206946v1/document
[63] Stiefenhofer, P., Deniz, C., 2025. The Role of AI in the Rise of the Ethical Consumer. In Proceedings of the SEPDER International Social Policies Congress, Sarajevo, Bosnia and Herzegovina, 17–18 May 2025.
[64] Stiefenhofer, P., Chen, Y., 2024. Industrial Artificial Intelligence: Stability of Cobb-Douglas Production Functions. Applied Mathematical Sciences. 18(4), 185–194.
[65] Stendere, K., 2020. The Latvian Tax Administration Has a Virtual System Assistant – Tom. IOTA. Available from: https://www.iota-tax.org/news/latvian-tax-administration-has-virtual-assistant-tom (cited 06 June 2024).
[66] Williams, C.C., 2021. Developing Efficient Risk Assessment System to Tackle Undeclared Work: A Toolkit. SSRN. DOI: https://dx.doi.org/10.2139/ssrn.3944128
[67] Law Gratis, 2025. Artificial Intelligence Law in Latvia. Available from: https://www.lawgratis.com/blog-detail/artificial-intelligence-law-at-latvia (cited 15 July 2025).
[68] Stiefenhofer, P., 2025. Zero-Labour Economy: The New Leisure Class. Journal of Chinese Economic and Business Studies. Special Edition: Navigating the Intersection of Digital Innovation and ESG: Pathways to Sustainable Development. (Submitted)
[69] Stiefenhofer, P., Deniz, C., Chen, Y., et al., 2024. The Future of Sustainable Finance: AI-Driven Sustainable Pairs Trading in Market-Neutral Investing. In: Walker, T., Gramlich, D., Sadati, A. (Eds.). Artificial Intelligence, Finance, and Sustainability. Palgrave Macmillan: Cham, Switzerland. pp. 111–142. DOI: https://doi.org/10.1007/978-3-031-66205-8_6
[70] European Commission, 2006. Risk Management Guide for Tax Administrations. Fiscalis Risk Analysis Project Group. Available from: https://taxation-customs.ec.europa.eu/system/files/2016-09/risk_management_guide_for_tax_administrations_en.pdf (cited 27 May 2024).
[71] Algorithm Watch, 2020. How Dutch Activists Got an Invasive Fraud Detection Algorithm Banned. Available from: https://algorithmwatch.org/en/syri-netherlands-algorithm/ (cited 27 May 2024).
[72] University of Amsterdam, 2023. The Dutch Childcare Benefit Scandal Underlines the Need for Explained AI Rules. Available from: https://www.uva.nl/en/shared-content/faculteiten/en/faculteit-der-rechtsgeleerdheid/news/2023/02/childcare-benefit-scandal-transparency.html (cited 26 August 2025).

[73] Government of the Netherlands, 2024. Dutch Government Presents Vision on Generative AI. Available from: https://www.government.nl/latest/news/2024/01/18/dutch-government-dutch-government-presents-vision-on-generative-ai#:~:text=The%20Dutch%20government%20has%20announced,via%20the%20Dutch%20AI%20Coalition (cited 27 May 2024).

[74] World Bank Group, 2022. Digital Transformation of Tax and Customs Administration. Available from: https://documents1.worldbank.org/curated/en/099448206302236597/pdf/IDU0e1ffd10c0c208047a30926c08259ec3064e4.pdf (cited 25 March 2024).

[75] Norwegian Ministry of Local Government and Modernisation, 2020. National Strategy for Artificial Intelligence. Available from: https://www.regjeringen.no/contentassets/1febbbb2c4fd4b7d92c67ddd353b6ae8/en-gb/pdfs/ki-strategi_en.pdf (cited 25 March 2024).

[76] European Commission, 2021. Norway AI Strategy Report. Available from: https://ai-watch.ec.europa.eu/countries/norway/norway-ai-strategy-report_en (cited 01 April 2024).

[77] Esayas, S.Y., Hauglid, M.K., 2024. Public Sector Digitalisation in Norway: Current Trends and Challenges in the Legal Framework. In: Motzfeldt, A. (Ed.). Public Digitalisation in a Legal Perspective: Status, Challenges and Opportunities for Nordic-Baltic Cooperation. pp. 193–230. Available from: https://munin.uit.no/bitstream/handle/10037/36470/article.pdf?sequence=3&isAllowed=y (cited 14 October 2025).

[78] Algorithm Watch, 2020. Pre-Crime at the Tax Office: How Poland Automated the Fight Against VAT Fraud. Available from: https://algorithmwatch.org/en/poland-stir-vat-fraud/ (cited 05 June 2024).

[79] Bal, A., 2022. Black-Box Models as a Tool to Fight VAT Fraud. In: Custers, B., Fosch-Villaronga, E. (Eds.). Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice. T.M.C. Asser Press: The Hague, Netherlands. pp. 225–236. DOI: https://doi.org/10.1007/978-94-6265-523-2_12

[80] OECD, 2020. Policy for the Development of Artificial Intelligence in Poland from 2020. Available from: https://wp.oecd.ai/app/uploads/2021/12/Poland_Policy_for_Artificial_Intelligence_Development_in_Poland_from_2020_2020.pdf (cited 05 June 2024).

[81] Dudkowiak, Putyra, 2025. AI Law in Poland. Available from: https://www.dudkowiak.com/ai-law/ (cited 15 July 2025).

[82] Olivares, B.D.O., 2020. Law and Artificial Intelligence in the Spanish Tax Administration: The Need for a Specific Regulation. European Review of Digital Administration & Law. 1(1–2), 227–234.

[83] Agencia Tributaria, 2025. Ethical Commitment in the Design and Use of Artificial Intelligence in the Tax Agency. Available from: https://sede.agenciatributaria.gob.es/static_files/Sede/Tema/Agencia_tributaria/Gobierno_abierto/compromiso%20_etico_uso_IA.pdf (cited 29 June 2025). (in Spanish)

[84] Erling, B., 2023. Automated Artificial Intelligence and Machine Learning Supported Tax Registration for Non-Swedish Customers. EIPA. Available from: https://www.eipa.eu/epsa/automated-artificial-intelligence-and-machine-learning-supported-tax-registration-for-non-swedish-customers/ (cited 01 April 2024).

[85] The Swedish Government, Ministry of Enterprise, 2018. Sweden's National Approach to Artificial Intelligence. Available from: https://dig.watch/resource/swedens-national-approach-to-artificial-intelligence (cited 29 June 2025).

[86] Ministry of Digital Transformation, Republic of Slovenia, 2023. On the Way to User-Centred Digital Transformation of Slovenia. Available from: https://www.gov.si/assets/ministrstva/MDP/On-the-Way-to-a-User-centered-Digital-Transformation-of-Slovenia.pdf (cited 18 July 2025).

[87] Republic of Slovenia, 2021. National Programme to Promote the Development and Use of Artificial Intelligence in the Republic of Slovenia by 2025. Available from: https://nio.gov.si/nio/asset/nacionalni+program+spodbujanja+razvoja+in+uporabe+umetne+inteligence+rs+do+leta+2025+npui (cited 07 June 2024).

[88] GOV.UK, Department for Science, Innovation & Technology, 2024. A Pro-Innovation Approach to AI Regulation: Government Response. Available from: https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response

[89] GOV.UK, 2024. "Criminals Should Be Aware" Says Minister as Government Upgrades AI Fraud Detection Tool. Available from: https://www.gov.uk/government/news/criminals-should-be-aware-says-minister-as-government-upgrades-ai-fraud-detection-tool (cited 5 June 2024).

[90] International Public Sector Fraud Forum, 2020. The Use of Artificial Intelligence to Combat Public Sector Fraud. Available from: https://assets.publishing.service.gov.uk/media/5e4545fe40f0b677be5fbd62/Artificial_intelligence_13_Feb.pdf (cited 05 June 2024).

[91] Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning, 2nd ed. Springer: New York, NY, USA. pp. 1–745.

[92] Ward Jr., J.H., 1963. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association. 58(301), 236–244. DOI: https://doi.org/10.1080/01621459.1963.10500845

[93] Kaufman, L., Rousseeuw, P.J., 2005. Finding Groups in Data: An Introduction to Cluster Analysis. Wiley: Hoboken, NJ, USA. pp. 1–342.
[94] Murtagh, F., Legendre, P., 2014. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? Journal of Classification. 31(3), 274–295. DOI: https://doi.org/10.1007/s00357-014-9161-z
[95] Floridi, L., Cowls, J., 2019. A Unified Framework of Five Principles for AI in Society. Harvard Data Science Review. 1(1), 2–15. DOI: https://doi.org/10.1162/99608f92.8cd550d1
[96] Dennett, D.C., 1997. Consciousness in Human and Robot Minds. In: Ito, M., Miyashita, Y., Rolls, E.T. (Eds.). Cognition, Computation, and Consciousness. Oxford University Press: Oxford, UK. pp. 17–29. DOI: https://doi.org/10.1037/10247-002
[97] Nissenbaum, H., 1996. Accountability in a Computerized Society. Science and Engineering Ethics. 2(1), 25–42. DOI: https://doi.org/10.1007/BF02639315
[98] Eubanks, V., 2018. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin's Press: New York, NY, USA. p. 260.
[99] Latour, B., 2005. Reassembling the Social: An Introduction to Actor-Network-Theory. Oxford University Press: Oxford, UK.