Digital Twins of Legislation for Explainable Automated Decision-Making in Administrative Law

Florian Schnitzhofer
Johannes Kepler University Linz
Linz, Austria
schnitzhofer@dke.uni-linz.ac.at

Anastasija Nikiforova
University of Tartu
Tartu, Estonia
anastasija.nikiforova@ut.ee

Christoph G. Schuetz
Johannes Kepler University Linz
Linz, Austria
schuetz@dke.uni-linz.ac.at

Abstract

Automated decision-making (ADM) in public administration must satisfy strict requirements of legality, transparency, and explainability. A promising but underexplored approach is the use of Digital Twins of legislation, which synchronize legal texts, semantics, and executable decision logic into computable representations. This paper investigates how such legislative digital twins can support explainable, rule-of-law–compliant ADM in administrative law. We conduct a grounded, inductive analysis of expert interviews and identify four feasibility conditions for trustworthy ADM: (1) selective automation of deterministic sub-decisions, (2) semantic standardization and ontological alignment, (3) computable legal structures linking natural language norms to machine-interpretable logic, and (4) governance and drafting adaptations enabling synchronized updates of law and code. Based on these findings, we derive design principles and propose a four-layer Digital Twin of Administrative Law (DTAL) architecture comprising statutory text, ontology, configuration, and executable logic. We illustrate the operation of the DTAL through a tourism contribution levy use case and compare DTAL outputs against human expert judgments and large language model baselines. The results show that DTAL enables deterministic, traceable, and reproducible decisions while LLM-based approaches remain inconsistent and non-transparent. This study contributes empirically grounded design principles, a reference architecture for computational law, and a mid-range process theory explaining when and how administrative norms can be operationalized through digital twins while preserving human oversight, legal traceability, and institutional legitimacy.

CCS Concepts

• **Computing methodologies** → **Artificial intelligence**; **Knowledge representation and reasoning**; • **Applied computing** → **Law**.

Keywords

Digital Twins of Legislation, Automated Decision-Making, Administrative Law, Law as Code

1 Introduction

Public administration is increasingly relying on *automated decision-making* (ADM) to improve efficiency, consistency, and accessibility [2]. However, ADM operates under strict constraints regarding the *rule of law*, and administrative decisions must remain legally correct, transparent, explainable, and reviewable [7]. Existing ADM systems often struggle to meet these requirements because the normative content of legislation, e.g., definitions, conditions, thresholds, and exceptions, remains embedded in natural-language text rather than in machine-interpretable structures [26]. This gap raises a question: *How can ADM in administrative law be improved while fully complying with the rule of law?*

Administrative decision-making is often routine, data-driven, and governed by formally articulated norms [5]. In practice, however, agencies frequently re-implement the same legal provisions across systems, leading to duplication, inconsistent outcomes, and costly maintenance when legal parameters change. These challenges point to a structural gap between natural language legislation and its operational execution in software.

A promising response to the challenges of ADM in administrative law is offered by digital twin principles. In engineering domains, digital twins provide continuously synchronized digital representations of physical systems, enabling traceability, consistency, and operational alignment [14]. Analogously, we explore whether such principles can be applied to legislation to bridge the gap between natural language legal texts and machine-interpretable representations. In this view, legislation is not only documented but also maintained as a computable artifact that can support consistent execution and explanation of administrative decisions. Thus, as a design response, we propose to centrally develop and publish digital twins of legislation, i.e., authoritative, machine-consumable counterparts to legal texts that serve as a single source of truth for computations and explanations [21, 23, 26].

We consider a *Digital Twin of Legislation* to be a synchronized, machine-interpretable representation of statutory and regulatory norms that encodes both semantic meaning and executable structure, serving as an authoritative basis for computation and explanation of legal outcomes. This concept enables a systematic coupling between legal text, formal semantics, and decision logic. Building on this idea, we propose an architecture for a *Digital Twin of Administrative Law* (DTAL) that supports the automation of deterministic, parameterizable sub-decisions while preserving human oversight, legal traceability, and institutional legitimacy. The design is grounded in the semantic, structural, and governance prerequisites required to make legislation computationally operational in a rule-of-law-compliant manner.

To empirically investigate the prerequisites for ADM in administrative law compliant with the rule of law, we conduct semi-structured interviews with senior experts in policy-making, electronic government, legal-technology development, legislative drafting, notarial practice, and legal informatics, using a grounded theory methodology [8]. From the interview data, we inductively derive four core constructs, namely automation scope, semantic alignment, computable structures, and governance capacity. These constructs form the basis for three propositions specifying the conditions

under which ADM can remain compliant with the rule of law. Importantly, the aim of digital twins of legislation is not to automate law-making or replace legal reasoning, but to support the safe automation of deterministic, parameterizable administrative subdecisions. Each outcome remains traceable to its originating legal provisions, ensuring transparency, auditability, and accountability.

We illustrate and validate the derived DTAL design principles using the Upper Austrian tourism contribution levy as a use case. Specifically, we compare the outputs of a manually implemented DTAL with those produced by direct LLM prompting, validating both approaches against a ground truth provided by a licensed Upper Austrian tax advisor. The comparison underscores the limitations of free-text reasoning in the absence of a structured legal model and motivates a DTAL approach that couples modern LLM chat interfaces with accurate, deterministic, and legally traceable ADM mechanisms.

The main contribution of this study is the definition of the concept of a DTAL along with the introduction of a DTAL architecture with four synchronized layers of text, ontology, configuration model, and executable logic that supports explainable ADM in administrative law. This study provides four empirically grounded design principles that translate the study's findings into actionable guidance, helping practitioners design and implement DTALs in a way that is robust, interoperable, and auditable in real administrative settings. Furthermore, this study offers a mid-range process theory explaining conditions under which administrative norms can be operationalized through digital twins.

The remainder of the paper is structured as follows. Section 2 reviews related work on computable law, digital twins, and legislative reform. Section 3 describes the employed research methodology. Section 4 reports the findings from the analysis of the interviews. Section 5 describes the design principles derived from the findings. Section 6 defines the DTAL concept and proposes a DTAL architecture. Section 7 demonstrates and evaluates the DTAL concept in the context of a use case. Section 8 provides a discussion and Section 9 concludes the paper.

2 Related Work

Logical models and argumentation frameworks articulate how normative conclusions are derived and justified [28]. From a public-administration perspective, Schartum [29] details the status quo pipeline for transforming legal texts into operational systems: The transformation begins with qualification of sources of natural-language normative text, the interpretation of those sources, and the specification of rules, then passes through formal analysis and representation (modeling, programming) to produce systems and algorithms, mainly by computer professionals. The resulting systems are tested and refined into a confirmed specification, with feedback loops between stages and "decision" points framing the transformation cycle. While this process view explains how legal texts are turned into code, the actual form of the representations that make those translations reliable and reusable is left open.

Accountable ADM in adminstrative law at scale requires shared, machine-readable models of legal knowledge. For this purpose, notable legal ontologies include FOLaw [35], which is a function-based legal ontology framework, LRI-Core [6], which conceptualizes five the "worlds" physical, mental, roles, abstractions, occurrences, and LKIF-Core [19], which employs multiple modules for knowledge interchange. By offering machine-readable semantic definitions for legal terminology, legal ontologies support automated reasoning and interoperability across applications [35]. In the public sector, the EU's Interactive Terminology for Europe (IATE) provides a standardized vocabulary, but deeper alignment is needed, e.g., anchoring definitions at a high legislative level so that subordinate regulations inherit the definitions, using language-neutral concept identifiers, and establishing persistent URIs for legal concepts [10]. Building on the representations, serialization standards and domain specific languages operationalize the models. LegalRuleML specifies XML-based encodings of norms that can be exchanged with rule engines [26], while Catala [23] offers a domain-specific language that compiles legislative logic with faithful links to the originating provisions, as shown in a French tax-code case study. At the process level, New Zealand's *Better Rules* initiative demonstrates how multidisciplinary co-drafting can surface inconsistencies early and produce executable prototypes alongside the text, closing the loop from ontology to implementation [4].

Once legal knowledge is extracted, legal automation, especially through *smart contracts*, translates machine-readable norms into deterministic execution environments. Smart contracts provide deterministic frameworks for automating contractual obligations and legal processes—even without the blockchain technology [17]. Using smart contracts, agreements can be automatically executed upon meeting predefined conditions, significantly reducing the need for intermediaries while enhancing trust and efficiency [1]. Thus, smart contracts demonstrate how legal agreements can be encoded as computational artifacts by combining the legal text, a configuration model, and the computational logic [1, 33], thereby complementing ontology-driven approaches with operational semantics.

To ensure that executable logic remains synchronized with authoritative texts, a promising direction is to create and publish *digital twins of legislation*—authoritative, machine-readable counterparts to legal texts that serve as a single source of truth [21, 31]. The concept of digital twins emerged in manufacturing for product life cycle optimization [14]. In this study, we follow the definition of Grieves and Vickers [15, p. 94], according to which "*the Digital Twin is a set of virtual information constructs that fully describes a potential or actual physical manufactured product from the micro atomic level to the macro geometrical level. At its optimum, any information that could be obtained from inspecting a physical manufactured product can be obtained from its Digital Twin.*" While digital twins are widely applied in aerospace, smart cities, and Industry 4.0, their use in legislative processes remains limited.

In summary, digital twins of legislation provide a foundation for ensuring that ADM logic remains systematically aligned with statutory sources under rule-of-law constraints. To this end, in this study we investigate the conditions (semantic, structural, and institutional) that need to be in place for ADM based on digital legislative twins to succeed.

Table 1: Pseudonymized List of Interviewees

ID	Role	Org. Type	Language
E1	Parliamentary Legislation Governance (EU)	Legislature (EU)	EN
E2	Parliamentary Policy, Advisor (EU)	Legislature (EU)	EN
E3	Legislative Counsel	National Parliament	DE
E4	Dean of Innovation	Education & Research	DE
E5	Notarial Practitioner	Legal Practice	DE
E6	ERP Vendor Lead (public sector)	Industry	DE
E7	Parliamentary IT Strategy (EU)	Legislature (EU)	EN
E8	Parliamentary Unit IT (EU)	Legislature (EU)	EN
E9	Parliamentary Innovation Lab (EU)	Legislature (EU)	EN

3 Methodology

To attain the set objective and develop of an architecture for a digital twins of administrative law (DTAL), we follow the design science research (DSR) methodology by Peffers et al. [27]. We also employ Straussian grounded theory [8] to empirically identify the core problem, gather requirements for solution objectives, and derive design guidelines at hand. To this end, we collect data through expert interviews and inductively analyze the interviewees' responses to uncover recurring themes and concepts. These inform the design of the DTAL architecture and evaluate this architecture using a realistic scenario. For this problem setting, we find a qualitative research design suitable because of the novelty of the phenomenon of "digital twins of legislation" in automated administration. An emergent, data-driven approach allows for theory building directly from expert experiences. Rather than starting from predetermined hypotheses, the research question is deliberately broad to allow for patterns to emerge from interview data through systematic coding and constant comparison.

3.1 Data Collection

We used purposive sampling to recruit participants with demonstrated extensive experience at the intersection of law, technology, and public administration [25] for semi-structured expert interviews. The experts include individuals from EU legislative bodies, a national parliament, public sector technology and policy roles, notary, and academia. Variation in geography, domain, and seniority was intentionally sought to enhance contextual richness. In total, nine experts were interviewed (Table 1). Each interview was conducted in person or via video call. Each session lasted between 60 and 90 minutes. Following a problem-centered approach, we began each interview with an open-ended question about the interviewee's experience with ADM or legal automation before following a flexible guide to probe topics if the topics did not emerge organically.

The key topics that were discussed included: (a) experiences translating laws into software (and pain points encountered), views on the feasibility of fully automating various types of laws, (b) opinions on the concept of publishing laws in machine-readable form, (c) technical and organizational prerequisites for ADM (e.g., data standards, IT infrastructure, skill sets), (d) and any changes anticipated or needed in legislative drafting practices to accommodate ADM. This approach allowed interviewees to narrate freely from their expertise while still covering our research objectives. Interviews were conducted in the participant's preferred language: German and English. All interviews were voluntary with gathered informed consent; recordings were made with permission and transcribed verbatim.

The qualitative research data, including a description of the study design, interview protocols, use case source-code, LLM prompts, and evaluation data, are available online [30].

3.2 Data Analysis and Rigor

Collected data were analyzed using Straussian grounded theory [8] coding techniques, following the iterative process of open coding, axial coding, and selective coding. Transcripts were coded in MAXQDA and in the open coding phase, we broke down the data into discrete concepts by assigning codes to meaningful segments of text (words, sentences, or paragraphs).

Next, we proceeded to axial coding, wherein we examined how these codes relate to each other and clustered them into higher-level categories and sub-categories. During axial coding, we wrote memos to capture our emerging understanding of each category and to record any propositions or hypotheses (for example, memos on the importance of explainability as an intervening condition for accepting ADM systems). Through axial coding, a set of prominent themes began to coalesce.

Finally, during selective coding, we refined resulting codes into a coherent set of core categories to answer our central research question. We focused on categories that were well-supported by data (appearing across multiple interviews) and were highly relevant to the research objectives. We selected, anonymized, and translated quotes before writing integrative memos to relate categories.

We determined that theoretical saturation was reached after conducting five interviews: No substantially new codes or concepts emerged in the last transcripts, and each core category was well-supported by multiple sources. We also took several steps to enhance the credibility and rigor of our analysis: we continuously compared codes and incidents both within and across interviews (constant comparison technique) to ensure consistency; we maintained an audit trail of codes, memos, and category development; and we conducted member checks by providing a summary of our preliminary findings to two interviewees to verify that our interpretations resonated with their experiences.

Grounded theory served as the ex-ante justification stage of our DSR by enabling systematic discovery of problem characteristics and design requirements for DTAL. The four grounded theory categories formed the empirically grounded objectives and boundary conditions for artifact design, consistent with the first two stages of the Peffers et al. [27] model. During design and development, these categories guided architectural decisions across DTAL's text, ontology, configuration, and logic layers, ensuring both relevance and rigor. An ex-post evaluation compares DTAL computation with expert-validated outcomes and LLM baselines. The results are reported in the evaluation section and used to assess utility,

correctness, and explainability in accordance with DSR evaluation criteria.

4 Findings

In this section, we summarize the main findings from the interviews, synthesizing the dominant themes that emerged from the expert interviews on *when* and *how* ADM can be introduced in administrative law in a way that remains legally acceptable and technically robust. Rather than advancing an "AI-first" narrative, participants emphasized concrete limits and design requirements: selective automation of deterministic sub-tasks, semantic standardization through ontological alignment, high-quality computable representations with legal traceability, and governance adaptations in legislative drafting. We therefore organize the results into four categories (F1–F4), which also provide the empirical basis for the subsequent design principles and the proposed digital-twin approach.

4.1 F1: Partial Transformation, Not Every Law can be Automated

A dominant theme among the experts was that only parts of the legal landscape can be feasibly transformed into deterministic algorithms, at least with current technology and within rule-of-law requirements. Not every law is automatable, and attempting to fully replace human legal reasoning with AI or code in all areas of law was seen as neither practical nor desirable. One interviewee bluntly stated that *"I cannot be satisfied with half-right solutions as a jurist. I don't want something that is 'probably' correct; I need to know* yes *or* no*"*, rejecting the notion that probabilistic AI systems (which offer, say, 95% accuracy) are acceptable for legal decisions. He added, *"I don't want AI [in such cases]; it is completely unsuitable, because AI is a probability calculation."* This reflects a commonly expressed insistence on certainty - legal decisions often require definitive true/false determinations (guilty or not guilty, eligible or not eligible), whereas machine learning algorithms typically provide probabilistic outputs.

Many experts stressed that stochastic *"black-box"* AI is a *"no-go"* in core legal determinations, as it violates the principle of explainability and predictability fundamental to the rule of law. In other words, if an algorithm cannot explain its reasoning or consistently produce the same output from the same input, it would not be trusted in a legal context. This sentiment was echoed across interviews – the notion of an inscrutable AI judge making decisions with opaque logic was described as *"inconceivable"* and something that *"would never gain approval of the courts"* (as one expert, himself a certified but not practicing judge, emphasized).

Participants distinguished traditional algorithms from machine learning. Formulaic calculations (e.g., taxes, standard fees) were viewed as suitable and largely accepted: *"We accept that every quarter the system automatically calculates our social insurance contributions … because it's formulaic."* Automated traffic fines were offered as another low-ambiguity example. By contrast, matters involving credibility assessment or open-textured terms (e.g., asylum) were seen as ill-suited: *"You can't decide that based on documents alone."* Discretion exists for a reason encoding it as fixed thresholds risks injustice and brittle over-fitting to edge cases.

In summary, the experts converged on the view that deterministic automation is appropriate only for clearly specified, low-discretion legal sub-tasks, whereas core legal determinations should remain under human authority, complemented by structured escalation mechanisms and explanations that are explicitly traceable to the applicable legal provisions.

4.2 F2: Semantic Standardization and Ontological Alignment

The second category centers on the need for consistent semantics in legislation to enable reliable automation. Experts repeatedly pointed out that one of the biggest hurdles in creating machine-interpretable laws is the ambiguity and inconsistency of legal language across different statutes, jurisdictions, and agencies. For an ADM system, it requires that all relevant terms and concepts are clearly defined and used uniformly—in other words, a shared ontology and vocabulary for the law. As one interviewee explained, when coding a law one must achieve semantic unification: *"You have to define terms unambiguously for this application field. But even defining terms unambiguously is not simple."* This quote highlights a dual insight: on the one hand, precise definitions are indispensable for computation (a computer cannot handle a concept that is intentionally vague or context-dependent) while on the other hand, drafting truly unambiguous definitions in law is itself a challenge, as natural language meaning can shift with context and interpretation.

One legal expert gave a telling example: *"In tax law, how long are you considered a child'? It makes a difference: in criminal law, when I'm no longer a child, I can be tried as an adult. In tax law, a 'child' might be defined for dependency deductions. These definitions can vary."* This underscores that even a seemingly straightforward term like "child" can have different legal definitions (e.g., age cutoffs) depending on the context and purpose of the law. An ADM system would need to know exactly which definition to apply in each context. Without semantic alignment, a digital twin of legislation could produce incorrect decisions simply due to divergent interpretations of a key term.

The experts thus emphasized a shared legal ontology and the involvement of legal ontologists in drafting these machine readable legal representations. EU terminology resources such as IATE (Interactive Terminology for Europe) were cited as helpful baselines, but participants argued for deeper alignment: (1) definitions anchored at higher legislative levels, so subordinate rules inherit them; (2) language-neutral identifiers to align multilingual texts; and (3) persistent URIs to ensure dereferenceable meaning. Pairing concept identifiers with the European Legislation Identifier (ELI) framework enables unambiguous cross-referencing and versioning across jurisdictions, as each fragment of EU law is assigned a unique identifier [10]. The upshot is straightforward: machine-readable law requires machine-readable clarity. Investment in official dictionaries, thesauri, and ontologies is foundational [26].

4.3 F3: Transforming Legislation into High-Quality ADM-Supporting Structures

The third category deals with the practical transformation of legal texts into forms that can directly support ADM – in other words, how to turn traditional legislation into a *"digital twin"* structure

that a computer can execute or at least parse unambiguously. The experts acknowledged that current legislation is not conceived in a machine-readable form: laws are drafted in prose, published as PDFs or webpages, and often contain complex cross-references and conditional clauses that are not readily translatable into code. To harness ADM, existing laws will need to be re-engineered into high-quality, structured representations without changing their legal meaning. This is a multifaceted challenge involving format, content, and process quality.

Current publication formats (PDF, natural language legal text) impede parsing and maintenance. One IT leader argued: *"We need to move away from old formats towards XML-based approaches that emphasize content reusability and common vocabularies."* Participants pointed to legal rule markup (e.g., LegalRuleML) and domain-specific languages (e.g., Catala) as promising pathways to encode operative logic with traceability to provisions [23, 26]. Interoperability emerged as another pillar: standardized data exchange, unique identifiers (persons, entities, assets), and connected registers are needed to obtain the inputs a twin requires.

Precision is recognized critical, i.e., where conditions and parameters are determinable, the statute—or an official annex—should make them explicit. Vague phrases such as *"net income"* thwart consistent automation; annexed decision tables or executable specifications can disambiguate without altering legal intent. Interviewees emphasized explainability as non-negotiable: systems should articulate reasons in legal terms; if ML components are used upstream (e.g., document classification), their output must be auditable and never the sole basis of a final decision.

As one expert summarized, the transformation of legislation into code is *"legislative craftsmanship"*, not just IT engineering.

4.4 F4: Adapting Drafting Processes and Governance for ADM

The final core category addresses the institutional and procedural changes required in the legislative domain to accommodate and govern the introduction of ADM via digital legislative twins. Even with advanced technology, our experts believe that without changes in how laws are drafted, reviewed, and maintained, progress will remain patchy. Simply put, law-making itself must evolve to produce *"ADM-ready"* legislation. This entails cultural shifts among drafters, new collaborative workflows, and supportive governance structures (including political will).

A recurrent observation was that legislative drafting is a conservative, traditions-preserving field. One interviewee, a veteran legislative counsel, described the mindset: *"Law-making is a conservative domain. Few show openness to innovation; I come with my imaginative proposals and they say, 'No, we have always done it this way.'"*

Yet, the rise of machine *"actors"* executing law pressures the system to pursue greater precision or adopt dual-format drafting to accommodate both humans and machines. In response, several jurisdictions have begun formulating digital-ready legislative principles. For instance, France's Digital Republic Act and New Zealand's Better Rules initiative exemplify a commitment to transparency and multidisciplinary co-drafting [4, 12].

A recurring source of friction is intentional ambiguity, often arising from political compromises that defer critical details to later interpretation. However, this practice significantly hinders automation; vague clauses necessitate either arbitrary implementation choices or frequent human intervention. Where automation is a goal, participants urged more concretized agreements, with configurable parameters replacing open-ended language. Governance implications include multidisciplinary review councils for digital readiness, official publication of code alongside delegated acts, pilot-first rollouts with evaluation, and clear accountability lines for algorithmic errors and redress.

Change management is critical, given perceived threats to professional roles. As one interviewee observed, *"People will block it because they would, in a way, abolish themselves."* Roles must therefore evolve toward oversight, exception handling, and continuous improvement.

5 Design Principles

Informed by the expert opinions, we transform the observed patterns into design principles (DPs) for DTAL that address the identified conditions. Specifically, we map each category to its corresponding design implication, DTAL design principle, and architectural layer in the proposed architecture.

The four grounded theory categories provide the empirical foundation for the DTAL artifact. Table 2 links these categories to the corresponding design principles and architectural layers, ensuring traceability from qualitative findings to artifact design within the DSR framework.

This alignment clarifies how qualitative findings directly shaped the DTAL architecture and design principles, satisfying DSR requirements for grounded artifact construction.

5.1 DP1: Selective Automation and Partial Transformation

Classify the legal corpus into parts that can be safely and meaningfully automated and parts that require human judgment. Our study reaffirms this pragmatism from an empirical perspective. In practice, this means developing criteria for determining DT suitability and then selecting target laws or processes for digital twinning. Likely candidates are those with deterministic logic, minimal discretion, and high volume (e.g., tax calculations, fee assessments, license expirations), as by real implementations like automatic tax filing and traffic fine systems. Conversely, processes involving personal assessment (e.g., asylum, complex permitting) should retain human decision-makers [3]. Current probabilistic AI approaches, often used to interpret legal texts, contradict the principles of the rule of law, as they produce likely-correct decisions rather than definitive ones [28].

5.2 DP2: Semantic and Ontological Clarity

Establish a shared legal ontology and standardized terminology as infrastructure. The need for semantic alignment found in our study is well supported by prior work in legal informatics. Researchers have long developed core legal ontologies and argued for semantic web technologies in law [19]. Our findings add evidence that practitioners see real value in these efforts. The mention of IATE

Table 2: Mapping of Grounded Theory Categories to DTAL Design Principles and DTAL Architectural Layers

Grounded Theory Category (Finding)	Design Implication	DTAL Layer(s)
F1: Partial transformation, not every law can be automated	DP1: Selective Automation and Partial Transformation	Logic
F2: Semantic standardization and ontological alignment	DP2: Semantic and Ontological Clarity	Ontology
F3: Transforming legislation into high-quality ADM-supporting structures	DP3: Computable Law Structures	All
F4: Adapting drafting processes and governance for ADM	DP4: Process and Governance Adaptation	All

and similar projects indicates a translational path from academic ontology work to practical terminology management. A direct implication is that governments should invest in maintaining official machine-readable vocabularies for legal terms (possibly extending systems like IATE with more granular definitions for national law) and require that new legislation references these where applicable. Additionally, adopting unique identifiers for legal concepts can enable unambiguous cross-referencing in digital twins. For instance, the European Legislation Identifier (ELI) [10] provides a framework for identifying legal provisions across jurisdictions – pairing ELI with ontologies could allow algorithms to retrieve the precise definition of a term or the current value of a parameter (like a tax rate) from an authoritative source. In sum, our results suggest that moving toward a *"semantic backbone"* for legislation is a concrete prerequisite for ADM.

5.3 DP3: Computable Law Structures

Transforming laws into structured, executable formats is the technological heart of the DTAL concept. We deliberately separate the ontology, the configuration model, and the executable logic. This modular separation improves maintainability, and facilitates the independent evolution of legal concepts, parameter settings, and enforcement rules. Several countries and organizations are experimenting with exactly what our experts envisaged: publishing legislation in dual forms – human-readable text and machine-consumable code. For example, the government of New Zealand's Better Rules project demonstrated a process for co-drafting legislation with a corresponding Python rule prototype [4]. They reported benefits like early error detection and clearer logic flows. A concrete step could be requiring that any new major piece of administrative legislation comes with a *"digital twin annex"* – for instance, an accompanying open-source code module or decision table that represents the operative logic.

5.4 DP4: Process and Governance Adaptation

The necessity of legislative process adaptation found in our study aligns with emerging discussions in legal and public administration scholarship. The legislative process traditionally yields natural-language statutes that are interpreted anew whenever they must be operationalized in software [29].

Our work suggests an iterative, multidisciplinary drafting approach with concurrent publication, in which drafters produce the natural-language legal text alongside a corresponding, formally structured ontology (see Fig 1). Similar to version-control practices in software engineering, every change to the legal text prompts a parallel revision of the digital representation [13].

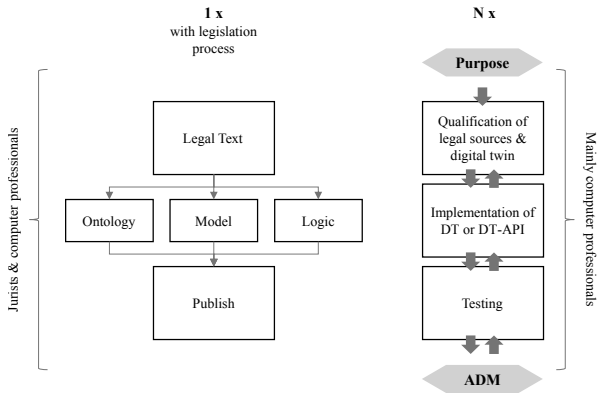

Figure 1: Proposed legislative process for publication of DTAL and implementation.

Not every statutory change necessitates a complete overhaul of the ontology and logic. For example, adjusting a tax rate only requires an update in the configuration-model layer. By contrast, more profound reforms—such as introducing a new levy category—necessitate modifying relevant classes or properties within the ontology and logic itself, thereby shifting the domain's conceptual boundaries. This distinction minimizes unnecessary complexity and helps maintain consistency across different instances of implementation, offering advantages for stakeholders who may otherwise have to interpret the text repeatedly [29]. This parallels the insertion of agile software development practices into the lawmaking process. Overcoming cultural resistance will require showing that such methods enhance legal quality rather than threaten it.

6 DTAL Architecture and Operationalization

By translating the experts' insights into design principles, we arrive at a conceptual structure for DTAL. The application of digital twins in administrative law builds on concepts traditionally associated with manufacturing and systems engineering, where the idea of creating a synchronized virtual model is well-established [14]. In these industrial contexts, digital twins serve to mirror physical systems or processes, thus facilitating real-time data capture, iterative improvements, and predictive analytics [24].

When ported into the legal domain, the same principles allow laws to be represented not merely as static texts but as digital models that continuously reflect legislative changes, practical implementations, and interpretive adjustments. Building on the digital-twin concept from engineering [15], we define a DTAL as follows.

Definition (Digital Twin of Administrative Law)**.** *A Digital Twin of Administrative Law (DTAL) is a continuously synchronized machine-interpretable representation of administrative law that encodes the normative content of statutes, regulations, delegated acts, and administrative guidance across the descriptive level (concepts, definitions, and relationships) and the decision-making level (eligibility conditions, parameterized computations, and decision outcomes). The DTAL enables automated or semi-automated computation of legally defined outcomes while preserving full traceability to the promulgated legal sources. When provided with complete and accurate case data, a DTAL yields the same legally relevant conclusion that would follow from applying the corresponding legal provisions. The DTAL also explains the conclusions through references to the governing norms.*

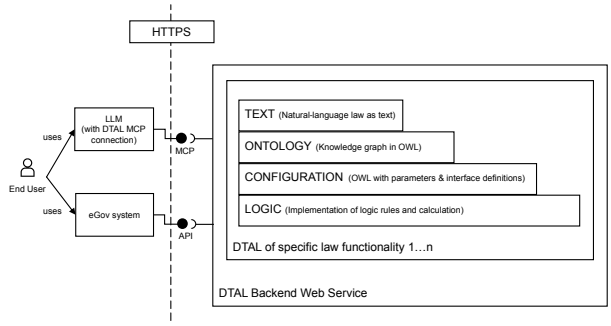

Figure 2: Software architecture of DTAL

Drawing on methods and design principles used for digital twins in manufacturing [15] as well as on blockchain-based smart contracts [1, 33], this research proposes a layered structure for a digital twin for administrative law comprising the following layers.

- **Statutory Text**: natural-language legal texts as traditionally published.
- **Ontology**: formal structure that defines (parts of) the law's semantics in a machine-readable way.
- **Configuration**: adjustable parameters in the axioms of the ontology (e.g., tax rates, exemption thresholds).
- **Executable Logic**: computational rules for execution of legal procedures (e.g., formulae, conditions).

The DTAL architecture (Fig. 2) is inspired by the principle of *separation of concerns*, which is central to smart contracts, which encourage modular design to allow incremental updates and transparent governance [17], with the proposed design principles applicable regardless of whether execution is deployed on a blockchain or on conventional infrastructure. While certain amendments to legislation only change numeric thresholds or percentage rates, others require foundational revisions to interpretive definitions. By distinguishing between ontological structures and configuration parameters, this framework aims to reflect legislative updates systematically, facilitating improved legal compliance and reduced duplication of coding efforts.

Similar to digital twins in manufacturing that evolve from individual components, e.g., single machines, to encompass entire supply chains [15], the legal domain can be digitized progressively. This incremental approach aligns with widely recognized best practices in both ontology engineering [34] and large-scale digital twin deployments [20], allowing pilot projects to refine proof-of-concepts before expanding to larger sets of statutes. Whereas industrial twins close their feedback loop through sensor coupling to a physical system [15], a DTAL closes its loop through the legislature: administrative practice, audit logs, and case outcomes expose drafting gaps or unintended effects, prompting legislative amendments that propagate as versioned updates to the affected DTAL layers.

A central building block of the DTAL architecture is its implemented legal logic: the executable rules that operationalize the underlying norms while preserving traceability from each computational outcome back to its legal grounds. In the DTAL, this traceability is not an afterthought but a design requirement. The logic layer is implemented in a high-level, general-purpose programming language (e.g., Java, Python, Prolog) with a logging functionality to explain the derived and calculated algorithmic results. Our use case prototype was implemented in Python (the source code is available online [30]), chosen for its flexibility, readability, and mature ecosystem of data-processing and API frameworks [4]. In our earlier prototype, we experimented with a Prolog-based implementation, which, although expressive in logic representation, proved difficult to interpret for both computer scientists and legal practitioners, but could nevertheless be a viable alternative.

The DTAL's logic layer exposes one or more services through standardized interfaces (typically RESTful APIs), making its computational functions accessible to external systems. These interfaces are also published via the Model Context Protocol (MCP), enabling direct integration with agentic AI applications. All API specifications and configuration metadata are documented within the configuration model to ensure traceability and transparency.

The central element in constructing the digital legal twin for administrative law is the ontology, which provides a formal semantic structure for describing key entities, relationships, and logical constraints within a statute [16]. In conventional legal informatics, ontologies have primarily been used for assisting with document analysis and advanced search [9]. While these ontologies are well-suited for legal reasoning and knowledge representation, they are not optimized for ensuring the Rule of Law in administrative law through automated decision-making. A fundamental distinction of the DTAL approach for administrative law is that the ontology does not merely serve as a taxonomic classification of legal concepts but rather as a dynamic, multi-layered semantic bridge connecting (1) legislative text, (2) real-world administrative data, and (3) the computational logic required for implementation.

7 Use Case: Upper-Austrian Tourism Contribution Levy

To demonstrate the proposed multi-layered architecture, we analyzed a concrete administrative scenario, namely the tourism levy implemented in one of Austria's federal provinces. Under the Oö. Tourismusgesetz [2018], approximately 29,000 businesses (e.g., hotels, restaurants, tour operators) must calculate and pay a yearly tourism contribution. The actual assessment is managed centrally by a state agency, but businesses are responsible for computing their own levy in their bookkeeping or via spreadsheets. This fragmented

implementation leads to an administrative burden on the businesses and a risk of inconsistent calculations or errors. It represents an opportunity for harmonization through a DTAL. Essentially, publishing the levy's calculation rules as a digital twin that all parties can use or integrate into their systems, ensuring consistency, transparency, and seamless interoperability across systems.

We evaluate DTAL as a DSR artifact by assessing its utility—i.e., *can it compute legally correct results?*—, correctness (alignment with expert-defined outcomes), and explainability (ability to justify outputs with legal references) following the DSR framework ([18]) evaluation logic. To this end, we implemented a DTAL prototype for the Upper Austrian tourism contribution levy following the proposed architectural layers. We encoded the official formulas and exemptions in an executable logic module. The translation of the statutory text in natural language into the four layers of DTAL is described in detail in [32]. We also developed an ontology of key concepts, such as what qualifies as "tourism turnover" and which exemptions apply. Finally, we defined configuration parameters that capture contextual variability, for example, the applicable rate per revenue bracket, which may change by year or by region.

The outputs produced by the DTAL were evaluated against two baselines. First, a licensed Upper Austrian tax advisor provided the legal ground truth for each scenario, establishing the normative reference for correctness. Second, we compared the DTAL against an LLM-only implementation of the same legal task, representing a contemporary AI-based alternative for translating natural language legislation into executable decisions. This comparative design allowed us to assess not only legal correctness, but also determinism, traceability, and robustness across different implementation paradigms.

For the evaluation of DTAL against the baselines, we compiled a dataset of 100 representative, synthetic scenario cases, covering diverse municipalities, sectors, and revenue levels. Each scenario's result was validated by a licensed Upper Austrian tax advisor, who provided the ground-truth payable amounts and corresponding legal justifications.

We ran the DTAL prototype in three iterations in comparison with two state-of-the-art LLM model versions, namely, OpenAI GPT-5 (Thinking) and OpenAI GPT-5 (Pro), on the same 100 cases and compared the results. The DTAL computed the exact payable amount and the results were traceable to the executed code. Both LLM models received identical prompts in three independent runs and generated plausible, but factually incorrect or inconsistent answers, therefore demonstrating the limitations of free-text reasoning without a structured legal model.

The evaluation also demonstrates how the DTAL artifact operationalizes the four design principles we defined earlier. DP1 (*Selective Automation*) is reflected in the correct computation of deterministic eligibility and calculation rules, while human judgment remains required for non-deterministic elements, confirming the intended boundary of automation. DP2 (*Semantic and Ontological Clarity*) is observable in the system's ability to resolve legal concepts consistently across statutory provisions, enabled by the shared ontology underlying the DTAL layers. DP3 (*Computable Legal Structures*) is validated through the execution of structured logic derived directly from statutory text, producing traceable outputs that LLM baselines fail to replicate reliably. Finally, DP4 (*Process and Governance Alignment*) is illustrated by DTAL's reliance on versioned statutory inputs and configuration parameters, ensuring that updates to the legal sources propagate transparently through the computational model.

Together, these observations (see Table 3) show that the DTAL artifact not only performs correctly, but also embodies the design principles required for explainable and legally robust ADM.

8 Discussion

Our study identified four interrelated conditions for responsible ADM in administrative law, namely *partial automation*, *semantic alignment*, *computable legal structures*, and *governance adaptation*. Each of these empirically grounded findings addresses a critical dimension of making automated decision-making explainable and rule-of-law compliant, and each resonates with challenges and insights from prior IS and public sector digitalization literature.

We distilled a set of design principles (DP1–DP4) that guide the development of rule-of-law-compliant ADM and propose a novel Digital Twin of Administrative Law (DTAL) architecture—a four-layer reference model comprising the legal text, a legal ontology, a configuration layer for parameters, and executable code serving as a blueprint for practitioners to develop explainable ADM systems that maintain a traceable link between software and the law.

Our findings suggest that a DTAL-based approach to automated decision-making can improve efficiency, explainability, and legal correctness in administrative settings. At the same time, the present study has limitations that affect the generalizability of these findings. Our qualitative sample is Europe-centric and composed of senior domain experts, which may underrepresent frontline administrative and citizen perspectives. We mitigated threats to internal validity threats through constant comparison, memoing, and member checks, but future research should examine DTAL across additional jurisdictions and include case workers, implementers, and affected citizens to further validate the organizational and societal applicability of the approach.

We suggest that future research explores cross-jurisdictional DTAL pilots, ontology standardization, LLM usage to generate DTALs, and governance models to ensure that digital legislative twins enhance fundamental principles of law. In view of the EU AI Act [11], which classifies most public-sector ADM as high-risk), our DTAL already follows an architecture designed to meet core obligations, i.e., risk management, technical documentation, logging, human oversight, and transparency. Future work will consider a detailed dedicated evaluation of DTAL in relation to the EU AI Act's conformity-assessment and post-market monitoring requirements.

This study contributes a mid-range process theory explaining *when* and *how* administrative norms can be operationalized through digital twins. The theory is grounded in four empirically derived constructs: (1) automation scope, (2) semantic alignment, (3) computability structures, and (4) drafting and governance capacity. We theorize that ADM feasibility is driven by the interaction of semantic clarity and computable structures, moderated by governance readiness, and bounded by the automation scope of the underlying norm. These relationships explain variance in the suitability of different administrative processes for automation and articulate the conditions under which DTAL-based computation maintains

Table 3: Comparative Evaluation of DTAL, Expert, and GPT-5 Model Outputs

Dimension	Expert (Ground Truth)	DTAL	GPT-5 (Thinking)	GPT-5 (Pro)
Method type	Manual (human expert)	Deterministic algorithmic (coded law)	Generative AI (LLM with step-by-step reasoning)	Generative AI (LLM, direct answer)
DP3: Accuracy (over iterations)	Ground Truth (baseline)	100% (algorithmically deterministic)	1. run: 0% 2. run: 83% 3. run: 33%	1. run: 24% 2. run: 26% 3. run: 33%
DP3: Legal traceability	Full (explicit legal refs)	Full (statute links)	Often missing / wrong refs	Often missing / wrong refs
DP2: Semantic correctness	Correct use of inputs	Inputs formally validated	Not consistent – misinterprets or ignores inputs	Not consistent – misinterprets or ignores inputs
DP3: Explanation quality	High (detailed, no errors)	Consistent, law-based; no hallucinations	Not consistent; prone to errors / hallucinations	Not consistent; prone to errors / hallucinations
DP4: Determinism & reproducibility	Consistent (manual analysis)	Yes (deterministic output)	No (stochastic output varies)	No (stochastic output varies)
DP1: Selective automation	Manual (human expert)	Selection done once manually with coding	n/a	n/a
DP4: Comparative performance	Manual interpretation (slow and recurring effort)	Instant (seconds)	LLM thinking process (seconds–minutes)	LLM thinking process (minutes)

rule-of-law guarantees. This theory extends IS research on digital public infrastructures by identifying semantic and institutional contingencies that shape digitalization outcomes in administrative law. Drawing on four emergent constructs, we formulate three propositions that together state the necessary conditions for rule of law compliant ADM.

- **Proposition 1.** *Semantic alignment is a necessary condition for credible ADM.* When legal concepts lack harmonized definitions and stable identifiers, computational representations cannot reliably reproduce the normative meaning required for lawful administrative decisions.
- **Proposition 2.** *Computable structures positively affect ADM feasibility only when governance capacity supports synchronized updating.* Machine-interpretable logic enhances reliability and transparency, but only if supported by legislative and administrative processes that ensure timely and authoritative updates to underlying legal norms.
- **Proposition 3.** *Automation is viable only for deterministic, parameterizable sub-decisions.* Normative elements requiring open-textured reasoning, proportionality assessment, or contextual interpretation must remain under human judgment, whereas deterministic eligibility tests and threshold computations can be automated without undermining legal legitimacy.

Taken together, these propositions define the semantic, structural, and institutional boundary conditions under which administrative norms can be operationalized through digital twins while preserving explainability, accountability, and compliance with the rule of law.

9 Conclusion

This paper demonstrates that digital twins of legislation can enhance ADM in the public sector when implemented within well-defined legal, semantic, and institutional boundaries. Through a grounded analysis of expert insights, we identified critical conditions under which ADM systems remain compliant with the rule of law. These insights informed the development of a framework for a *Digital Twin of Administrative Law* (DTAL), which bridges the gap between natural-language statutes and executable code. We developed a mid-range process theory explaining how administrative norms can be operationalized via digital twins, encapsulated in four key factors (automation scope, semantic alignment, computable structures, and governance capacity). From this theory, we derived three propositions that specify the semantic, structural, and institutional conditions required for ADM to uphold fundamental legal principles. Thus, our work provides both a conceptual blueprint and an empirical rationale for deploying hybrid ADM models that combine deterministic computations with human oversight.

By integrating empirical findings with a design artifact, we establish a foundation for trustworthy, transparent, and efficient digital administration. Public agencies and lawmakers can build on these results to pursue digital transformation initiatives that do not sacrifice explainability or legal integrity, moving closer to the vision of a digitally enabled yet accountable government.

References

[1] Jason Grant Allen. 2022. *Smart Legal Contracts: Computable Law in Theory and Practice.* Oxford University Press.

[2] Kevin D Ashley. 2017. *Artificial intelligence and legal analytics: new tools for law practice in the digital age.* Cambridge University Press.

[3] Aaron Baird and Likoebe M Maruping. 2021. The next generation of research on IS use: A theoretical framework of delegation to and from agentic IS artifacts. *MIS quarterly* 45, 1 (2021), 315–341.

[4] Tom Barraclough, Hamish Fraser, and Curtis Barnes. 2021. Legislation as code for New Zealand.

[5] Mark Bovens and Stavros Zouridis. 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 (2002), 174–184.

[6] Joost Breuker, Andre Valente, Radboud Winkels, et al. 1997. Legal ontologies: a functional view. In *Procs. of 1st legout workshop on legal ontologies.* 23–36.

[7] Danielle Keats Citron and Frank Pasquale. 2014. The scored society: Due process for automated predictions. *Wash. L. Rev.* 89 (2014), 1.

[8] Juliet Corbin and Anselm Strauss. 2014. *Basics of qualitative research: Techniques and procedures for developing grounded theory.* Sage publications.

[9] Cleyton Mario de Oliveira Rodrigues, Frederico Luiz Gonçalves de Freitas, Emanoel Francisco Spósito Barreiros, Ryan Ribeiro de Azevedo, and Adauto Trigueiro de Almeida Filho. 2019. Legal ontologies over time: A systematic mapping study. *Expert Systems with Applications* 130 (2019), 12–30.

[10] European Legislation Identifier (ELI) Task Force. 2017. European legislation identifier (ELI): Technical specifications. Publications Office of the European Union. https://eur-lex.europa.eu/eli-register/about.html

[11] European Union. 2024. Regulation (EU) 2024/1689 of the European Parliament and of the Council. Regulation (EU)..

[12] French Republic. 2016. Digital Republic Act (Law No. 2016-1321 of 7 October 2016). https://www.wipo.int/wipolex/en/legislation/details/16380

[13] Asunción Gómez-Pérez, Mariano Fernández-López, and Oscar Corcho. 2004. *Ontological Engineering: With Examples from the Areas of Knowledge Management, E-Commerce and the Semantic Web.* Springer.

[14] Michael Grieves. 2014. Digital twin: manufacturing excellence through virtual factory replication. *White paper* 1, 2014 (2014), 1–7.

[15] Michael Grieves and John Vickers. 2016. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary perspectives on complex systems: New findings and approaches.* Springer, 85–113.

[16] Thomas R Gruber. 1993. A translation approach to portable ontology specifications. *Knowledge acquisition* 5, 2 (1993), 199–220.

[17] James Hazard and Helena Haapio. 2017. Wise contracts: smart contracts that work for people and machines. In *Trends and communities of legal informatics. Proceedings of the 20th international legal informatics symposium IRIS.* 425–432.

[18] Alan R Hevner, Salvatore T March, Jinsoo Park, and Sudha Ram. 2004. Design science in information systems research. *MIS quarterly* (2004), 75–105.

[19] Rinke Hoekstra, Joost Breuker, Marcello Di Bello, Alexander Boer, et al. 2007. The lkif core ontology of basic legal concepts. *LOAIT* 321 (2007), 43–63.

[20] Werner Kritzinger, Matthias Karner, Georg Traar, Jan Henjes, and Wilfried Sihn. 2018. Digital Twin in manufacturing: A categorical literature review and classification. *Ifac-PapersOnline* 51, 11 (2018), 1016–1022.

[21] Anita Lamprecht. 2025. Digital Twins of Law: Embracing Complexity. In *Liquid Legal–Sustaining the Rule of Law: Artificial Intelligence, E-Justice, and the Cloud.* Springer, 115–135.

[22] Land Oberösterreich. 2018. Oö. Tourismusgesetz 2018 (LGBl. Nr. 3/2018). https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=LrOO&Gesetzesnummer=20000953 Retrieved Feb. 1, 2025..

[23] Denis Merigoux, Nicolas Chataing, and Jonathan Protzenko. 2021. Catala: a programming language for the law. *Proceedings of the ACM on Programming Languages* 5, ICFP (2021), 1–29.

[24] Elisa Negri, Luca Fumagalli, and Marco Macchi. 2017. A review of the roles of digital twin in CPS-based production systems. *Procedia manufacturing* 11 (2017), 939–948.

[25] Lawrence A Palinkas, Sarah M Horwitz, Carla A Green, Jennifer P Wisdom, Naihua Duan, and Kimberly Hoagwood. 2015. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and policy in mental health and mental health services research* 42, 5 (2015), 533–544.

[26] Monica Palmirani, Guido Governatori, Antonino Rotolo, Said Tabet, Harold Boley, and Adrian Paschke. 2011. LegalRuleML: XML-based Rules and Norms. In *Rule-Based Modeling and Computing on the Semantic Web: 5th International Symposium, RuleML 2011 – America (Lecture Notes in Computer Science, Vol. 7018).* Springer, 298–312. doi:10.1007/978-3-642-24908-2_30

[27] Ken Peffers, Tuure Tuunanen, Marcus A Rothenberger, and Samir Chatterjee. 2007. A design science research methodology for information systems research. *Journal of management information systems* 24, 3 (2007), 45–77.

[28] Henry Prakken and Giovanni Sartor. 1997. *Logical models of legal argumentation.* Springer.

[29] Dag Wiese Schartum. 2020. *From Legal Sources to Programming Code: Automatic Individual Decisions in Public Administration and Computers under the Rule of Law.* Cambridge University Press, 301–336.

[30] Florian Schnitzhofer, Anastasija Nikiforova, and Christoph G. Schuetz. 2026. Qualitative Research Data for "Digital Twins of Legislation for Explainable Automated Decision-Making in Administrative Law"'. doi:10.5281/zenodo.18366112

[31] Florian Schnitzhofer and Christoph Schütz. 2025. Towards a scalable Architecture for Legal-Ontologies integrated into Digital Twins of Administrative Law. In *SEMANTICS 2025 Developers Workshop.*

[32] Florian Schnitzhofer and Christoph G. Schütz. 2025. Towards Translating Natural Language Normative Text into a Digital Twin of Administrative Law. In *Proceedings of the International Workshop on Translating Natural Legal Language into Formal Representation (NLL2FR 2025)*, Ken Satoh, Georg Borges, Hannes Westermann, and May Myo Zin (Eds.). Turin, Italy, 157–163.

[33] The Accord Project. 2024. Accord project concepts. https://docs.accordproject.org/docs/accordproject-concepts.html Retrieved Feb. 1, 2025..

[34] Mike Uschold and Michael Gruninger. 1996. Ontologies: Principles, methods and applications. *The knowledge engineering review* 11, 2 (1996), 93–136.

[35] André Valente. 1995. *Legal knowledge engineering: A modelling approach.* IOS Press.