

Review of doctoral dissertation

Yannis Charalabidis
Professor of Digital Governance
Director, Digital Governance Research Centre
University of the Aegean, Greece

Author of the Dissertation: Jaromir Durkiewicz

Title: Digital Government Value Logic - Conceptual and Benchmarking Model

Discipline: Management and Quality Sciences

Supervisor: Tomasz Janowski

Summary of the evaluation

The dissertation addresses a genuinely important and under-disciplined problem in the digital government (DG) field: how country-level digital government benchmarking (DGB) can be aligned with the logic of public value generation, rather than functioning primarily as a static ranking exercise. The author argues, rather convincingly in my view, that many internationally recognized benchmarking initiatives (European Commission, United Nations, OECD, World Bank) provide “context-free snapshots” of selected aspects of DG and, while useful as descriptive instruments, often underperform as actionable diagnostic tools for policy and management. The core scientific move of the thesis is therefore to introduce a framework-Digital Government Value Logic (DGVL)-that conceptualizes DG as a technology-enabled value-generating system composed of interacting conceptual components and value-generating processes among them, and then to show how this conceptualization can inform a more policy-relevant benchmarking practice.

The dissertation is organized around four research objectives (RO1-RO4), spanning: (i) integrating scattered ideas in the literature into DGVL, (ii) positioning DGVL as an extension of current benchmarking approaches, (iii) providing a methodological guide for designing DGVL-based benchmarking instruments, and (iv) demonstrating feasibility through a pilot application. The empirical/operational part is expressed as a DGVL-based benchmarking model (DGVLB), developed in four studies (Project Study, Concept Study, Method Study, Pilot Study).

Methodologically, the work combines a qualitative comparative analysis of major DGB projects using an Activity Theory-inspired lens, Formal Concept Analysis (FCA) to extract “conceptual packages” of indicators and a Data Envelopment Analysis (DEA)-based operationalization of value-generation efficiency, extended with projected targets, reference sets, best-practice matrices, and cross-efficiency computations to mitigate limitations of DEA self-appraisal.

The pilot implementation demonstrates the model on five representative value-generating processes (F2T, D2S, T2GP, S2U, C2D), mapped to public value categories and supported by relevant literature.

In overall terms, this is a coherent and ambitious dissertation that aims to move DGB from “measuring and ranking” towards bench learning and actionable diagnosis, while remaining realistic about the constraints imposed by available international datasets. The work is strongest where it:

- (a) identifies the conceptual mismatch between value-generation narratives in DG research and the typical mechanics of DGB instruments,
- (b) provides a structured method pipeline from concept to measurement (project selection, concept extraction, method toolset and pilot), and
- (c) presents tangible analytical outputs (efficiency scores, reference peers, targets, best-practice matrices) that, if adopted carefully, could indeed inform decision-making.

There are also limitations that should be made explicit in the evaluation: the heavy reliance on available composite indicator systems creates a risk of “benchmarking the benchmarkers”; some choices (e.g., elimination of UN-EGS and WU-IDGR, the selection of processes, or the data harmonization strategy)

are defensible but would benefit from clearer sensitivity discussion; and the pilot remains an intentionally limited demonstration rather than a full-scale DG performance assessment, which the author openly acknowledges.

Despite these reservations, I find that the dissertation demonstrates:

- Theoretical knowledge in DG conceptualization, public value debate, and benchmarking systems
- Ability to conduct independent research, including concept development, methodological integration, and the construction of a multi-source benchmarking dataset
- An original contribution in framing DGB through explicit value-generation logic, providing a replicable methodological guide and pilot instrument.
- An overall coherence in aligning objectives, research and validation of the research results.

On that basis, and consistent with the criteria stated in the institutional instructions for doctoral dissertation reviews, I provide below a detailed assessment and conclude with a clear recommendation.

1. Scientific novelty and contribution to the discipline

1.1 Originality of the research problem

The thesis identifies a gap that is both practical and theoretical: many benchmarking projects provide comparative measurement but do not explicitly operationalize how digital government generates public value, and therefore struggle to produce action-guiding insights (as opposed to descriptive league tables). This gap is not merely rhetorical. It becomes operationally consequential when governments use benchmarking outputs to justify investments, design reforms, or communicate progress. Without an explicit logic linking components (legal/organizational foundations, technology platforms, data governance, service quality, citizen engagement) to outcomes (service usage, governance quality, democratic strength), benchmarking can easily devolve into “counting what is countable” rather than “measuring what matters”.

The author’s key conceptual claim is that the value-generation imperative should be placed at the core of DGB design: benchmarking should more literally behave like benchmarking-supporting learning from peers, diagnosing inefficiencies, and mapping improvement pathways—rather than producing static rankings. This is an important repositioning that aligns with contemporary policy thinking about evidence-informed governance, performance management, and capability maturity—while remaining grounded in the DG literature’s public value tradition.

1.2. The DGVL conceptual framework as a contribution

The DGVL framework is presented as a technology-enabled system for generating public value through value-generating processes among conceptual components. This formulation is valuable because it explicitly unifies two ways DG is frequently described:

- DG as a process (stages moving from inputs to outcomes), and
- DG as a system (interacting components and actors).

The dissertation claims conceptual kinship to earlier frameworks such as the “DG value chain” (Heeks) or public value-oriented evaluation approaches (e.g., Loukis), while asserting originality in the *shape and formulation* of DGVL and in its explicit readiness for measurement translation. I find this positioning reasonable: the novelty is not that “public value matters” (that is established), but that the author provides a structured logic model intended to guide benchmarking instrument design, including the identification of value-generation points and value-generation processes, and the interpretation of observed “inefficiencies” in DG value generation.

1.2 The DGVLB methodological guide as a contribution

Beyond conceptualization, the thesis contributes a “how-to” pathway: the DGVL-based benchmarking model (DGVLB) and its staged development across four studies. This is a meaningful contribution because much DG scholarship remains either conceptual without operationalization, or empirical

without sufficiently explicit conceptual scaffolding. Here, the author attempts to bridge that gap by specifying:

- how to select and evaluate existing DGB projects (Project Study),
- how to extract interpretable “conceptual packages” from indicator sets via FCA (Concept Study),
- how to operationalize value-generation efficiency via DEA and related diagnostics (Method Study),
- and how to instantiate this in a pilot analysis of recognized DG problems expressed as processes between components (Pilot Study).

This staged approach is, in my view, one of the thesis’s strongest features: it provides a replicable research design that other scholars can adapt or critique constructively.

2. Definition and approach to objectives

The dissertation defines its core aim as introducing a conceptual and methodological framework positioning DGVL as a foundation for more relevant DGB, and demonstrating practical usefulness. The four research objectives (RO1-RO4) are stated clearly in Table 1.1 and are logically sequenced: integration to extension to method guide to feasibility demonstration.

A notable strength is that the objectives are not treated as purely declarative: the structure of the dissertation and the staged DGVLB development operationalize them. In particular:

- RO1 (integration of scattered ideas) is addressed by the conceptual framework development and literature synthesis linking DG, public value, and benchmarking critiques.
- RO2 (extension of current DGB approaches) is realized through the critique of conventional instruments and the argument for value-generation alignment, including the decision to treat existing instruments as “intermediate rather than final products.”
- RO3 (methodological guide/toolset) is addressed through the method pipeline: data sourcing, concept packaging, DEA specification, cross-efficiency, and output diagnostics.
- RO4 (feasibility and capacity to address substantial issues) is addressed via the pilot analysis of five processes mapped to policy/management problems.

The thesis also states its broader positioning as use-inspired basic research (“Pasteur quadrant”) and describes its reasoning as abductive, where findings from one stage shape subsequent conceptual positions. This is an appropriate philosophical stance for a dissertation that aims to build a conceptual-methodological instrument while grounding it in empirical feasibility and real-world benchmarking constraints.

3. Theoretical framework and background

3.1. Digital government definitions and conceptual clarity

The dissertation correctly notes the definitional plurality in DG and provides an explicit adopted definition aligned with public value creation (not merely digitization of service delivery). This matters because a value-generation logic requires clarity on what counts as “DG” and what counts as “value.” The thesis positions e-government and digital government as related but distinct, citing the OECD framing that DG is the use of digital technologies as part of modernization strategies to create public value.

This conceptual grounding is adequate, and I particularly appreciate that the author addresses terminological consistency (using “Digital Government” even when citing “E-Government” sources, for clarity). That said, any conceptual framework that aims at measurement translation must remain vigilant against definitional drift: the more components and processes are introduced, the more important it is to preserve internal coherence between conceptual meaning and indicator operationalization.

3.2. Public value as the normative core

The dissertation anchors the framework in public value literature (citing, among others, the inventory perspective and the DG public value discourse). The pilot process selection references the public value catalogue of Twizeyimana & Andersson (2019), mapping processes to values such as improved public services, administrative efficiency, trust, open government capabilities, democracy, and social well-being. This is a sensible approach: it makes public value explicit rather than implied.

One theoretical challenge-common in public value research-is that public value can become an umbrella term so broad that it risks being analytically diffuse. The thesis mitigates this by operationalizing value generation through process efficiencies between components and outcomes, rather than claiming to “measure public value” directly in an abstract sense. The inclusion of IMPACT indicators (via B-SGI governance/democracy-related measures) is a practical attempt to approximate higher-level outcomes.

3.3. Benchmarking theory and critique of existing instruments

The thesis’s diagnosis-that major DGB projects often provide static cross-sectional snapshots, are not explicitly value-generation oriented, and therefore provide limited actionable feedback-is well motivated. It is also consistent with broader evaluation and performance management critiques: indicators often reflect what is easy to collect, institutions inherit legacy designs, and ranking incentives can distort measurement priorities.

The author’s decision in Study 1 to examine five major DGB initiatives (EU-EGB, UN-EGS, OECD-DGI, WB-GTMI, WU-IDGR) is well chosen for relevance, and the elimination of UN-EGS and WU-IDGR from the model-development pathway is justified through criteria of conceptual coherence, technical realization, and EU relevance. (I discuss below how this decision might be strengthened through explicit sensitivity argumentation or external validity discussion.)

3.4. DGVL as synthesis of process logic and system logic

The conceptual move to unify process logic (technology-enabled operations) and system logic (interactions among components) under the value-generation lens is theoretically sound and-crucially-useful for benchmarking design. The author explicitly cites that some prior conceptualizations integrate these perspectives (e.g., DG value chain, DG value activity system), and positions DGVL in that family while aiming for a more measurement-ready conceptual platform.

4. Methodology and research design

The dissertation adopts mixed methods (qualitative + quantitative) and abductive reasoning in a stage-based process where each stage informs the next. This is well aligned with the nature of the research problem: a purely quantitative study would struggle to justify indicator selection and conceptual structure; a purely qualitative study would struggle to demonstrate feasibility and analytical value.

4.1 Study 1: Project Study

In Study 1, the author conducts a qualitative analysis of five DGB projects and applies criteria around conceptual coherence, technical realization, and EU relevance. This is a strong design choice: DGVLB explicitly depends on reusing existing instruments as data sources, so the quality and suitability of those instruments must be assessed upfront.

A methodological point worth highlighting is that this evaluation is not purely descriptive; it functions as a gating mechanism that shapes what “DG components” can be operationalized later. This is both a strength (pragmatism, feasibility) and a limitation (instrument dependence). The author is transparent about this trade-off.

4.2 Study 2: Concept Study

Applying FCA to extract “conceptual packages” from indicator sets is a creative and defensible methodological decision. FCA has the advantage of being mathematically grounded for discovering formal concepts from tagged attributes, and it can help avoid purely intuitive clustering. The outcome-five packages (foundation, technology, data, service, citizen orientation)-is interpretable and aligns with common DG component categories.

However, FCA outcomes depend heavily on the quality of tagging and the initial design of the indicator-attribute mapping. The dissertation includes supporting tables (keywords extraction, tags coverage, reduced context, retained concepts) in Chapter 4. For the purposes of doctoral-level rigor, this is appropriate, yet I would encourage the author to make the epistemic status of the packages explicit: they are formal concepts derived from the selected indicator universe, not necessarily “true” ontological partitions of digital government. This is not a flaw; it is simply the correct interpretation. Doing so strengthens the scientific honesty and prevents overclaiming.

4.3 Study 3: Method Study

The operational core of DGVLB is the use of DEA to model the “relative technical efficiency” of transforming inputs into outputs within DG value-generation processes. DEA is suitable when (i) multiple inputs/outputs exist, (ii) the goal is relative efficiency, and (iii) functional forms are unknown or not assumed. The thesis uses standard variants (CCR output-oriented; BCC input- and output-oriented), and explains why CCR input-oriented is omitted (yielding identical scores to output-oriented CCR).

A particularly good methodological decision is the inclusion of **cross-efficiency** to mitigate DEA’s self-appraisal bias and improve discriminatory power when many DMUs appear efficient under their own chosen weights. The author computes cross-efficiency using a neutral (“arbitrary”) method rather than benevolent/aggressive schemes, which is defensible for a benchmarking instrument aimed at neutrality.

Moreover, the dissertation does not stop at efficiency scores; it generates a richer diagnostic package: projected target values, reference sets, best-practice matrices. This is essential because efficiency scores alone can be too abstract for policy use. The best-practice matrix for F2T is a concrete example of how peers can be interpreted as mentors, and at what “rate” inefficient units might learn from efficient ones.

4.4 Study 4: Pilot Study

The pilot study is explicitly framed as a demonstration, not as a complete evaluation of EU DG. This is scientifically appropriate; it avoids the common trap of “pilot results overinterpretation.”

The pilot uses multiple data sources with partial timing mismatch and coverage challenges: EU-EGB (2021-2022 average), OECD-DGI (2023 results based on 2020-2022 collection), WB-GTMI (2022 update), B-SGI impact (2022 issue), Eurostat usage (2021). The dissertation explains these constraints and documents them in Table 4.15. The data harmonization includes converting values into a 0-100 scale and handling missing OECD-DGI values for Germany and Greece by median-difference adjustment between editions.

Process selection is based on (i) recognition in literature, (ii) match to public value categories, (iii) assignable input-output components. The five processes selected are:

- FOUNDATION to TECHNOLOGY (F2T)
- DATA to SERVICE (D2S)
- TECHNOLOGY to GOVERNANCE/POLICY (T2GP)
- SERVICE to USAGE (S2U)
- CITIZEN to DEMOCRACY (C2D)

This set is conceptually coherent and spans from enabling conditions to higher-order outcomes. Importantly, it also reveals where the benchmarking instrument might face the hardest attribution problems (especially T2GP and C2D), a point the thesis should treat with extra caution (see Section 5 of this review).

5. Relevance of results, discussion, and conclusions

5.1 What the pilot results show

The results section provides DEA efficiency scores by country and process, comparing CCR-O, BCC-I, and BCC-O variants. It also provides cross-efficiency summaries with process-wise and average cross-efficiency values, which is helpful for interpretation.

The outputs are appropriate for the thesis purpose: demonstrating feasibility, generating diagnostic artifacts, and illustrating how a value-generation lens changes what is being “benchmarked.” The results show, for example, that a country can be efficient in one process and relatively inefficient in another, which is a key argument against one-dimensional rankings.

However, and this is crucial for interpretive integrity: DEA efficiency is not causal impact. It is a relative position against an empirically observed frontier given the chosen indicators and model assumptions. The author seems aware of this general issue, and the pilot framing helps. Still, the dissertation would benefit from consistently policing language around “effects” and “influence,” especially in higher-level processes (technology to governance; citizen mechanisms to democracy), where confounding factors are substantial and outcomes are multi-determined.

5.2 Limitations and boundary conditions

The dissertation mentions data shortages and limited range. I would additionally emphasize several boundary conditions that are intrinsic to this kind of work:

1. Indicator endogeneity and circularity: benchmarking instruments often include proxies that reflect policy choices, reporting capacity, or survey design rather than underlying DG capabilities. DGVLB inherits these limitations, even as it reorganizes indicators into a value-generation logic.
2. Temporal misalignment: combining datasets from 2020-2022+ introduces risks if fast-moving domains (digital participation, platform maturity) shift quickly. The author documents timing constraints; the interpretive caution should match that transparency.
3. Frontier sensitivity: DEA frontiers can be sensitive to variable selection and sample composition. The cross-efficiency extension helps, but the thesis could further strengthen credibility by showing limited sensitivity checks (e.g., dropping one indicator package, altering scaling assumptions, or comparing with simple regression-based diagnostics).
4. Institutional and contextual drivers: value generation depends on administrative culture, legal tradition, fiscal capacity, political stability, and societal trust. DGVLB can flag inefficiency patterns but should avoid implying that the measured DG components alone “cause” governance outcomes.

To the author’s credit, the dissertation’s key claim is not that it provides causal identification; rather, it provides a **value-generation aligned benchmarking logic and toolset** that produces more actionable comparative feedback. Interpreted under that promise, the results are relevant and the limitations manageable.

6. Formal aspects

The dissertation is written in English and includes appropriate academic structure: introduction, conceptual/theoretical grounding, methodological development, staged studies, validity/evaluation chapter, discussion and implications, and appendices for datasets and matrices.

The inclusion of tables and appendices (datasets, best practice matrices) is particularly appropriate given the methodological nature of the work. Referencing appears extensive and consistent with standard academic practice (based on the visible citations throughout the excerpted sections). The thesis demonstrates attention to terminological clarity and definitional anchoring.

If I were to indicate areas for formal improvement, they would be largely “presentation polish” rather than substantive:

- ensure consistent use of key terms (e.g., DG vs DGB vs DGVL vs DGVLB) in chapter transitions;

- add short “chapter end summaries” that restate what was accomplished relative to RO1-RO4, to help readers track the argumentative spine;

None of these issues constitute a formal deficiency; they are recommendations for strengthening readability and defensibility.

7. Conclusion and recommendation

In view of the above assessment, I conclude that the dissertation Digital Government Value Logic - Conceptual and Benchmarking Model by Jaromir Durkiewicz meets the formal and substantive requirements expected of a doctoral dissertation, as articulated in the applicable framework referenced in the institutional review preparation instructions.

The work presents a coherent and original conceptual-methodological contribution to the discipline, demonstrates the candidate’s general theoretical knowledge and capacity for independent research, and offers a feasible, well-motivated approach to improving the policy relevance of digital government benchmarking.

Therefore, I submit a positive conclusion and request: Admission to defend (positive review).

8. Questions for the candidate

To complement the review and support a substantive discussion, I propose the following questions:

1. **On conceptual boundaries:** DGVL is designed as a “technology-enabled system for generating public value.” What are the boundary rules for the system? For example, where do you place broader institutional trust, media freedom, or fiscal capacity-inside the DG system, in its environment, or as moderating conditions?
2. **On causal language and interpretation:** How do you propose that policymakers should interpret an “inefficient” score in a process like TECHNOLOGY to GOVERNANCE/POLICY, given the many confounding influences on governance quality? What interpretive safeguards do you recommend?
3. **On DEA model choice and frontier effects:** Why did you choose the combination of CCR-O and BCC variants as the core comparison set, and how should a policymaker interpret differences between CCR and BCC efficiency classifications for the same process?
4. **On “benchlearning” operationalization:** The best practice matrices imply a mentor-mentee learning structure. What is your proposed real-world mechanism for translating these matrices into actionable learning programs (e.g., peer reviews, targeted capability exchanges, EU-level governance arrangements)?

Samos, Greece

15th March 2026

Yannis Charalabidis