Towards Synthetic and Balanced Digital Government Benchmarking

Jaromir Durkiewicz
Department of Applied Informatics in Management
Gdańsk University of Technology, Poland
jaromir.durkiewicz@pg.edu.pl

Tomasz Janowski
Gdańsk University of Technology, Poland
Danube University Krems, Austria
tomasz.janowski@pg.edu.pl

Abstract

Reliable benchmarking is essential for effective management of the government digitalization efforts. Existing benchmarking instruments generally fail to support this target. One problem is the diversity of instruments, resulting in a split image of digital progress and adding ambiguity to policy decisions. Another problem is disconnect in assessing progress between digital and traditional "analog" governance, lending support to a dangerous idea that countries can compensate for lack of progress in their governance systems by simply digitalizing them. This paper provides a path to addressing both problems by: aggregating relevant indicators of the World Economic Forum's Network Readiness Index (NRI) to obtain a single synthetic measure of digital government, balancing this measure with progress in analog governance using World Bank's Worldwide Governance Indicators (WGI), calculating new measures for the latest editions of NRI and WGI, and discussing results. Technically, the paper applies multidimensional linear ordering and factor analysis.

1. Introduction

When treating digital government as a deliberate, perceptible, designable and measurable arrangement of tools, decisions and actions, there is a rudimentary need for data-based benchmarking that is able to capture and compare digital government performance among countries, states, cities and other entities. Established instruments such as the NRI or the United Nations' e-Government Development Index serve different goals, implement different models and target different users. However, those who rely on such instruments to perform consequential decisions, i.e. policymakers and public managers responsible for overseeing the construction of digital government, encounter at least two issues here: 1) the diversity of benchmarks makes the image of digital government progress split and ambiguous, and thus difficult to translate to concrete policy decisions; and 2) the

"digital" measures do not say much about the impact of digitalization on traditional "analog" governance, thus lending support to a dangerous idea that countries can compensate lack progress in their governance systems by simply digitalizing them.

This situation is confusing for policy makers who introduce relevant policies; for public managers who implement such policies; for citizens who need a clear picture of how digitalization will change their interactions with authorities; and for businesses that need assurances of good governance and the rule of law. Such expectations may be compromised by nefarious state actors claiming that because their country is digitally successful, it is also well governed. Such claims can be supported by arbitrary selections of benchmarks to create an ambiguous and partitioned picture of the reality.

This paper provides a path to addressing both problems. First, the identified problems are expressed through qualitative statements and mapped into numbers and correlations. Second, a single synthetic indicator is built to measure digital government by integrating relevant NRI indicators. This step applies factor analysis to identify two uncorrelated dimensions of digital government and construct one synthetic indicator using those dimensions. Third, we integrate this new measure with WGI variables that represent analog governance. Using multidimensional linear ordering, this step creates a combined indicator that expresses progress on both (digital and analog) sides of governance. It should be stated that these new measures are not new benchmarks in terms of frameworks, data collection, etc., but the result of calculation carried out upon existing indicators.

Concerning terminology, we treat "analog" governance as synonymous with public governance. The term does not have a settled definition [1], e.g. [2] associates governance with "structures and decision-making processes that allow a state … to conduct affairs", and [3] defines it as "the exercise of … authority … to manage a nations affairs". Here, we are interested in those aspects of governance that can be associated with state policy and public value development [4]. Although we chose not to subscribe

to any particular conceptualization, we treat WGI as a conceptual and quantitative basis for representing "analog" governance. On the digital side, we adopted the conceptualization provided by NRI, and discuss the meaning of digital government in Section 2.2.

This work studies conceptual and practical issues that hinder digital government assessment, if treated as a subject connected to traditional governance, in contrast to existing measurements that consider this subject in isolation, disconnected from its "analog" foundations. Hence this work targets the needs of: 1) policy-makers searching for balanced paths to government digitalization that improves public governance; 2) practitioners seeking reliable and versatile methods of measuring digital government; and 3) researchers looking for a quantitative grasp of digital government studied in a conceptual manner.

The rest of the paper is organized as follows. Section 2 examines digital government benchmarking and identifies related problems and conceptual issues. Section 3 presents the research problem, and methods and tools used to address it. Section 4 describes the results achieved, followed by discussion in Section 5. Section 6 provides the summary and limitations of this work and outlines directions for future work.

2. Background and Problem Formulation

The aim of this section is to present theoretical concepts and practical problems associated with digital government benchmarking. The outcome relies on the literature review, performed on Scopus using the search term ("digital government" OR "e-government" OR "e-governance") AND ("theory" OR "model" OR "framework") AND ("evaluation" OR "benchmark" OR "measurement"), and using the snowballing technique. We examined the papers to confirm their relevance. Theoretical and practical aspects of digital government benchmarking are outlined in Section 2.1, and two problems addressed – diversity and digital-analog disconnect of existing digital government benchmark instruments – are covered in Sections 2.2 and 2.3 respectively.

2.1. Digital government benchmarking

Benchmarking is a relatively new form of activity, [5], dating back to late 1980s, although its historical roots can be found much earlier [6]. With digital government arriving in early 1990s, and the first digital government benchmarking instruments appearing in early 2000s, lack of universal consensus of what and how should be measured to establish the maturity of digital government is unsurprising.

Various efforts have been underway to equip digital government evaluation with solid theoretical foundations, e.g. in terms of the "readiness" construct [7][8], using "socio-technical models" to underpin key design elements of information systems [9], structuring the benchmarking activity using Activity Theory [10], employing advanced multi-criteria decision support models [11], and introducing context-driven benchmarking for digital government services [12]. The reasons, scope and methods of digital government benchmarking were also provided as recommendations for practitioners [13].

The wealth of approaches to digital government benchmarking give rise to measurement instruments that vary in their goals, examined units, technical construction, etc. For instance, [14] lists seven instruments that capture country-level data, from global studies like the United Nations' e-Government Survey or the World Economic Forum's (WEF's) NRI, to local studies such as McKinsey's study on Middle East digital economy; [15] presents a collection of benchmarks divided into academic and commercial instruments; and [16] introduces indices that confirm diversification of available instruments. According to [17], various approaches towards the methods and goals of measurement are reflected by digital strategies of individual countries.

The literature also includes critique of existing benchmarks, taking different positions, covering various aspects, proposing different improvements. For example, [7] questions the data-gathering approach and limited informational value of the UN benchmark, [13] points out limitations of benchmarks focusing on adoption and use rather than outcomes and impact, and [18] uncovers various technical limitation of benchmarks including their failure to "differentiate between static websites and highly integrated and interactive portals".

Our study does not focus on the construction of yet another benchmarking instrument, but on the question of how to effectively use the information provided by existing instruments and how to merge such information with external information to create a new value. This standpoint can be associated with various statements from the literature. For instance, the statement that "any ranking system needs a final single scale" [15] is refuted by the diversity of instruments, [19] notices "forgetting citizens' needs, demands or expectations" in the domain of digital government, while [20] observes the failure of instruments "to capture the expected transformative effects of ICT on government". Hence the central premise underlying our study is using existing indicators to uncover a consistent message about the actual effects that digitalization has on governance.

2.2. Problem 1: Diversity of benchmarks

The benchmarks vary substantially, e.g. in terms of their logic. From a practical viewpoint, while some specialization is needed when accompanied by interpretative care [15], the diversity and mutual incoherence of benchmarks is confusing, especially to policy makers that expect a clear assessment of their digital policies. The following literature-based observations explain the causes of such diversity.

The first cause is the capacity and negotiability of the concept itself. Defining digital government as government use of digital technology trivializes the problem, claiming that harnessing a new facility creates a new entity as a matter of course, while focusing on two umbrella terms – government and digital – leaves room for interpretation. Definitions vary in scope – from information to democracy, technology – from computers to Internet, and subject – from citizens to stakeholders [21]. Over time, expectations grow towards transformation rather than information or service supply [22]. Technological and transformative structure and operation are at the core of the concept [23]. Besides such variety, digital government is also considered to undergo evolution [24]. In summary, the richness of the concept implies heterogeneity of its measurement approaches.

The second is the assumption of benchmark's coherence with its own objectives and framework rather than the digital government landscape. This is not necessarily negative – [15] associates it with targeting "to answer specific and narrow questions", but comes with "tunnel effects" and "risk avoidance" due to deep specialization [25]. Similarly, limited representation of certain categories draws upon "restricted definition of eGovernment" [26]. Nonetheless, established international benchmarks manifest their goals rather clearly, e.g. with NRI grouping numerous indicators into categories. Attempts at unifying benchmarks via synthetic models [27] or composite indices [28] are also noted.

2.3. Problem 2: Digital-analog disconnect

While varying in focus, logic, etc. the benchmarks measure digital government in isolation, not part of the traditional "analog" government. However, this context is compelling, if not central, for evaluating digital efforts. Pure digital benchmarking can be still useful, but one should maintain some reserve when drawing conclusions on it. Some countries treat high positions in digital rankings as a target in itself [5], which can be stimulating for digital architects but confusing for those who seek reliable assessment of digitally-enabled governance in such countries.

The influence of digitalization on governance is expressed in declarations, normative statements and even mounted in digital government definitions. UN underlines its strategic role in human development [29]. OECD points at efficiency improvements, enhanced service quality and policy outcomes [30]. ITU suggests economic benefits – cost reduction, improved management and better procurement [31]. Benefits covering social issues, democracy, natural resources, education, etc. are also put forward [21].

Research literature brings various proofs of the transformative potential of digital government for better governance [32], transparency and structural change [33], civic involvement in democratic reform [34], and reshaping democratic governance [35]. Expectations are substantiated towards, e.g. citizen empowerment [36], support to democracy [37] and transformation and accountability [33]. In summary, expectations towards "analog" yields of digital government are strongly justified, thus assessing the "analog" impact of digital government is relevant.

Various data-based studies were carried out to examine the presence and strength of the digital-analog link. For instance, [38] studied digital-analog correlations via relevant measures on both sides, confirming conformity in the area of government effectiveness and uncovering discrepancy in the area of democracy. A link between digital government and good governance [39] was also examined, with digital technology causing 57% progress in good governance indicators [40]. The connection between WGI [41] and e-government benefits was conceptualized in [42], and confirmed strongly for the Balkan states [43] and weakly for Indonesia [44].

The analog-digital connection was also explored in the digital democracy domain. Confronting the UN's e-participation index with measures of actual democracy and Internet freedom, [45] concluded the failure of the index to deliver consistent results, offering elaboration of the e-participation framework [46]. A new e-Democracy index was offered to ease such dissonance [47]. According to [48], political regime and government capacity influence digital government. Also, positive impact on government effectiveness and civil liberties, but lack of such impact on the level of corruption were proven in [49].

The review above confirms interest in assessing digital government outcomes. Considering measures of outcome, most initiatives pay limited attention to the effects of digital transformation. For instance, NRI [50] measures user satisfaction, but cares little about the impact on policy and governance. In our view, only a comparison of benchmark's results with indicators external to it can bring a reliable assessment of digital government efforts.

3. Methodology

This section presents the approach adopted in this paper to address the two identified problems, problem 1 in Section 3.1 and problem 2 in Section 3.2. In each section, the problem is quantified, and a matching statistical technique is provided.

3.1. Problem 1: Diversity of benchmarks

Diversity of benchmarks could be regarded as a problem of data overflow, if different benchmarks raise the level of knowledge about digital government or merely create cognitive confusion. And even if the informational value is revealed, is this value unique or another expression of a more general message?

To illustrate this issue, we explored the NRI data, a comprehensive instrument that takes a broad view over digitalization issues, including governance [51]. The official dataset [52] contains data collected for 151 entities, mainly independent states, covering several years. In the analysis, we used the data for the most recent year, 2016. Among calculated indices we identified five that refer to digital government, three belonging to the "Government usage" pillar: 8.01) Importance of ICT to government vision, 8.02) Government Online Service Index, and 8.03) Government success in ICT promotion, and two to the "Social impacts": 10.03) ICT use and government efficiency and 10.04) E-Participation Index [50]. Table 1 describes this set in more detail, along with the respective operational labels used in our research.

Table 1. NRI digital government indicators, source: [52]

NRI indicator	Description	Label
8.01 Importance of ICTs to government vision	To what extent does the government have a clear implementation plan for utilizing ICTs to improve your country's overall competitiveness? [1 = not at all—there is no plan; 7 = to a great extent—there is a clear plan]	D_VISION
8.02 Government Online Service Index	The Government Online Service Index assesses the quality of government's delivery of online services on a 0-to-1 (best) scale.	D_SERVICE
8.03 Government success in ICT promotion	In your country, how successful is the government in promoting the use of ICTs? [1 = not successful at all; 7 = extremely successful]	D_PROMOTION
10.03 ICT use and government efficiency	In your country, to what extent does the use of ICTs by the government improve the quality of government services to the population? [1 = not at all; 7 = to a great extent]	D_EFFICIENCY
10.04 E-Participation Index	The E-Participation Index assesses, on a 0-to-1 (best) scale, the quality, relevance, and usefulness of government websites in providing online information and participatory tools and services to their citizens.	D_PARTICIPATION

The indices 8.01, 8.03 and 10.03 are based on the WEF's own surveys while 8.02 and 10.04 are based on the UN's E-Government Survey. The former are in the range 1-7 and the latter in the range 0-1. To ensure the same scale, we rescaled the UN indicators to the 1-7 range. After elimination of rows with missing data, 137 individual observations remained.

The correlation matrix, i.e. correlation coefficient value in the scale from negative 1 to positive 1, for the five examined variables is shown in Figure 1.

Figure 1. NRI variables – correlation matrix

All pairs are positively correlated, with four pairs revealing very strong correlation (above 0.9). This confirms the assumption of data redundancy – significant parts of the original information is shared across variables, and could be expressed with little information loss using fewer variables. The desirable outcome would be replacing fragmentary measures with a one that synthesizes the original information.

In search of the synthetic indicator, we decided to apply the dimensionality reduction technique, i.e. replacing a set of variables with a smaller set while saving most of the original information [53][54][55]. One of common approaches to dimensionality reduction are principal component (PCA) and factor analysis (FA) [56][57]. To ensure mathematical correctness of this approach, we confirmed that the PCA/FA assumptions are satisfied for our dataset.

To build the synthetic indicator, we followed the four steps described in [56][57][58]. The first checks the correlation of data as only correlated variables can benefit from this approach. As observed before, the NRI indicators are mostly strongly correlated. The second determines how many components (factors) extracted with PCA analysis are sufficient. We followed formal criteria, e.g. the Kaiser criterion

to retain the factors with eigenvalue bigger or equal to 1 [56][57][58]. The third step involves rotating the coordinate system axes to reflect original variables on the retaining factors in terms of "loadings". Among various algorithms, we chose varimax rotation [59]. The fourth step constructs the synthetic indicator on the basis of the outcomes of FA, after rotation.

Given that the total information, i.e. variation, from the original dataset is distributed among calculated components ordered by their variation share [56], this method is likely to determine a small number of components to represent the examined concept, i.e. digital government according to NRI. Hence, despite reducing the number of variables, the amount of information is left almost intact.

We found the approach effective, transparent – there are no arbitrary variable weighting assumptions or "hidden" algorithms, and relevant – it addresses the problem. However, the method has also its limitations [56]. Primarily, since the calculation is based on historical data, conclusions are applicable to a posteriori analysis, but there is no guarantee that the validity of such calculations will be preserved in the future. While statistical inference may address this limitations to some extent, this is beyond this study.

3.2. Problem 2: Digital-analog disconnect

To address lack of substantial digital-analog connection, we need a set of indicators to represent the performance of "analog" governance. To this end, we decided to use the World Bank's WGI, released on the basis of the arguably most comprehensive and enduring governance framework. The instrument measures six dimensions expressed in composite indicators [60]. The official dataset [61] contains data collected for over 200 countries captured on a yearly basis from 1996 to 2017, measured on a standardized scale with the mean 0 and standard deviation 1. To make this data comparable with the "digital" dataset,

we selected the matching 137 observations. Table 2 describes this set, along with the labels used.

The digital-analog disconnect problem can be also illustrated by means of correlation. The correlation matrix between variables belonging to the digital and analog datasets is depicted in Figure 2. The figure reveals that while some pairs are strongly correlated, the correlation is moderate to weak for most pairs. For example, the analog voice is poorly reflected by any digital indicator. The message is clear: there is no justification for extrapolating digital government progress with progress in "analog" governance. Thus pure digital measures should be treated with caution or, as below, should be augmented with analog components to yield a more balanced measure.

To this end, we decided to harness a multivariate technique called linear ordering [62]. The technique classifies research objects, e.g. countries, in regard to some synthetic latent measure that balances the relevant aspects expressed by directly measured input indicators [62][63]. The result is a one-dimensional ranking of development within a certain domain, in our case balanced digital-analog governance.
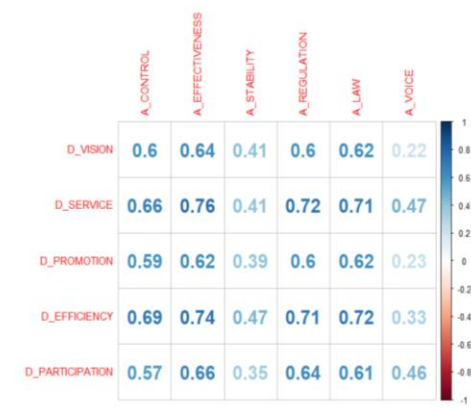

Figure 2. NRI vs WGI correlations

Table 2. Worldwide Governance Indicators, source: [61]

WGI indicator	Description	Label
Control of Corruption	Captures perceptions of the extent to which public power is exercised for private gain, including corruption, as well as "capture" of the state by elites and private interests.	A_CONTROL
Government Effectiveness	Captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.	A_EFFECTIVENESS
Political Stability and Absence of Violence	Measures perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism.	A_STABILITY
Regulatory Quality	Captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development.	A_REGULATION
Rule of Law	Captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.	A_LAW
Voice and Accountability	Captures perceptions of the extent to which a citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media.	A_VOICE

A common approach is to determine for each of the input variables whether it contributes to the phenomenon in a positive, negative or neutral way – in our case such contributions should be positive, and optionally how significant is the contribution [62]. Among existing ordering algorithms [62] we decided to use the Ideal Type, which identifies two theoretical objects, the best "ideal type" with maximum levels of indicators and "anti-ideal type" with minimum levels. For each examined object, a synthetic measure is then calculated comparing this object's performance to the distance between the ideal and anti-ideal types.

The key message is that this method allows for creating one combined digital-analog indicator, where performance on the digital side is balanced by performance on the analog side. This construction is in line with our key assumption that digital government is not in competition with, neither an alternative to "traditional" governance.

The above calculations were performed with MS Excel and R, including "psych" library for PCA/FA and one of the authors' own linear ordering program.

4. Findings

This section presents the results achieved by means of the approach introduced in Section 3.

First, we checked the digital variables against the PCA/FA assumptions. There was one outlier, United Arab Emirates with D_PROMOTION at 6.21. To prevent it from distorting the analysis, this case was excluded from factor calculations. Given 5 input variables, the remaining 136 records were deemed sufficient. Measure of the sampling adequacy, based on the Kaiser-Meyer-Olkin test – 0.73 and empirical p-value in Bartlett's test – 3.5e-224 confirmed that the dataset was adequate. Hence, we applied the synthetic indicator construction algorithm.

To determine the number of factors, the initial set of non-rotated PCA-extracted components was calculated. Table 3 shows the components ordered by their eigenvalues, i.e. the share of original variance. As components 1 and 2 together cover 97% of the original information, we decided that the optimal target number of factors for further analysis was two.

Table 3. NRI dataset - PCA-extracted components

	Component number				
	1	2	3	4	5
Eigenvalue	3.84	1.02	0.08	0.04	0.03
Cumulative share of variance	77%	97%	99%	99%	100%

Table 4, two varimax-rotated factors with normalized variable loadings, shows a legible arrangement of the factors. The main observation is that vision, promotion and efficiency – the NRI's variables measuring citizen perception – are mapped mostly to factor 1, with respective loadings of 0.95, 0.96 and 0.90, while service and participation – the variables "borrowed" from the UN E-Government Survey – are mostly mapped to factor 2, with respective loadings of 0.93 and 0.96. Hence the first factor could be identified with "citizen perception", while the other with "measured effectiveness" of digital government. Considering factor scorings for countries, obtained by summing raw scores [64] and scaling them to 1-7, the best factor 1 scores were: United Arab Emirates (6.15), Singapore (6.09), Qatar (5.72), and Rwanda (5.53), while the best factor 2 scores were: Singapore (6.45), Republic of Korea (6.36), United Arab Emirates (6.16) and the Netherlands (6.16).

Table 4. Variables and factors after rotation

Variable	Loadings on		Common variance
	factor 1	factor 2	
D_VISION	0.95	0.26	97%
D_SERVICE	0.33	0.93	98%
D_PROMOTION	0.96	0.25	98%
D_EFFICIENCY	0.90	0.39	96%
D_PARTICIPATION	0.23	0.96	98%
Sum of squared loadings:	2.79	2.07	n/a
Share in whole variance:	56%	41%	n/a
Share in explained variance:	57%	43%	n/a

The groupings identified two "composite" indicators, the first containing vision (weight 0.95), promotion (0.96) and efficiency (0.90), and the second service (0.93) and participation (0.96). In line with the previous procedure, further calculations resulted in the final synthetic indicator's (D_GOV's) weights for all five variables: vision – 0.20, service – 0.21, promotion – 0.20, efficiency – 0.18, and participation – 0.22, briefly summarized in Table 5.

Table 5. Weights of the synthetic indicator (D_GOV)

Variable	Original loading	Intermediate normalized weight (square factor loadings scaled to unit)	D_GOV weight
1st composite (high loadings on factor 1) – 57% variance			
D_VISION	0.95	0.34	**0.20**
D_PROMOTION	0.96	0.35	**0.20**
D_EFFICIENCY	0.90	0.31	**0.18**
2nd composite (high loadings on factor 2) – 43% variance			
D_SERVICE	0.93	0.48	**0.21**
D_PARTICIPATION	0.96	0.52	**0.22**

This way, we obtained statistics-proven weights to aggregate the original variables into one synthetic measure of digital government – D_GOV. The same scale applies, from 1 (worst) to 7 (best).

Table 6 presents 20 (among 137) best performing countries according to D_GOV. The list is topped by Asian and particularly Middle Eastern countries. Strong presence of well-developed Western countries and one South American country – Uruguay – is also noted. However, according to the 2016 Democracy Index [65], less than half of the top digital performers are "full democracies", with the rest classified as "flawed democracies" or "authoritarian regimes".

To balance the digital and analog measures, we applied the linear ordering Ideal Type algorithm to the variable D_GOV and the attributes of analog governance. To represent the latter, we applied six dimensions of WGI. Given that one synthetic digital indicator was put against six analog indicators, we gave D_GOV six times bigger weight. All indicators were standardized prior to analysis. The resulting measure of balanced digital-analog performance is called DA_GOV. For comparability with D_GOV, the value of DA_GOV was rescaled from 0-1 to 1-7.

Table 6 presents 20 (among 137) best performing countries according to DA_GOV. The DA_GOV ranking is more nuanced that D_GOV. While Singapore stays on top, Western democracies such as the United Kingdom, United States, Sweden, Switzerland, etc. achieved very high positions thanks to the solid digital and traditional governance performance. Notably, Uruguay is also at the top.

In two rightmost columns, Table 6 also presents the countries with the largest gains and the largest losses from considering their analog performance in addition to digital performance. Notably, African and generally poorer countries experienced the largest gains, e.g. Zambia's advanced by 72 positions. Another remarkable change are large gains for former socialist countries – Slovenia, Slovakia, Vietnam, Serbia, Ukraine, Poland, Romania and Tajikistan – suggesting that improvements in governance may overpass digital progress. On the other hand, Asian countries experienced the largest losses, e.g. Azerbaijan retracted by 59 positions.

In general, D_GOV and DA_GOV are strongly correlated – both Pearson and Spearman coefficients equal 0.80. This hints that while individual changes are noticeable, the range of analog corrections is not dramatic. However, it should be also noticed that D_GOV supplies half of the DA_GOV data content, thus such accordance is expected.

To make comparisons more systematic, we calculated the averages of D_GOV and DA_GOV regarding three country profiles: geographic – based on the UN' classification [66], political – based on the Freedom House's assessment of political rights and civil liberties [67], and economic – based on the World Bank's assignments to income groups [68]. The outcome is depicted in Table 7.

Table 6. D_GOV and DA_GOV - 20 best performances, 2016

No	Digital ranking		Balanced ranking		Analog advantage		Analog disadvantage	
	Country	D_GOV	Country	DA_GOV	Country	Change	Country	Change
1	Singapore	6.26	United Kingdom	6.39	Zambia	+72	Azerbaijan	-59
2	United Arab Emirates	6.15	United States	6.28	Trinidad and Tobago	+60	Bahrain	-50
3	Republic of Korea	5.89	Singapore	5.91	South Africa	+59	Bangladesh	-49
4	Bahrain	5.68	Uruguay	5.86	Slovenia	+57	China	-48
5	United Kingdom	5.65	Sweden	5.69	Slovakia	+47	Egypt	-48
6	Netherlands	5.65	United Arab Emirates	5.68	Uganda	+46	Colombia	-47
7	Japan	5.60	Switzerland	5.32	Swaziland	+45	Albania	-47
8	France	5.54	Spain	5.19	Vietnam	+44	Armenia	-46
9	Estonia	5.54	Netherlands	5.15	Switzerland	+43	Kazakhstan	-44
10	United States	5.52	New Zealand	5.07	Seychelles	+38	Ethiopia	-40
11	Qatar	5.42	Norway	5.04	Tanzania	+38	Argentina	-40
12	Israel	5.38	Portugal	4.90	Zimbabwe	+38	Ecuador	-38
13	New Zealand	5.34	Qatar	4.82	Serbia	+37	Kenya	-36
14	Australia	5.30	Sri Lanka	4.71	Ukraine	+37	Brazil	-30
15	Canada	5.22	Luxembourg	4.70	Poland	+36	Gambia	-30
16	Norway	5.21	Japan	4.67	Paraguay	+36	Côte d'Ivoire	-29
17	Malaysia	5.21	Republic of Korea	4.65	Tunisia	+33	Bolivia	-29
18	Saudi Arabia	5.20	Slovakia	4.61	Romania	+33	Iran	-27
19	Luxembourg	5.18	Vietnam	4.60	Thailand	+30	Honduras	-26
20	Uruguay	5.16	Slovenia	4.54	Tajikistan	+29	Costa Rica	-24

Table 7. D_GOV and DA_GOV – average scores for geographical, political and economic groupings

Subset	D_GOV	DA_GOV
Geographic		
Africa	3.25	3.16
Americas	3.81	3.54
Asia	4.25	3.73
Europe	4.35	4.17
Oceania	5.32	4.73
Political		
Free	4.35	4.13
Partly free	3.55	3.32
Not free	3.78	3.32
Economic		
High income	4.79	4.51
Upper-middle income	3.79	3.41
Lower-middle income	3.49	3.24
Low income	3.01	2.96

On average, the best performing group, both in terms of D_GOV and DA_GOV, is Oceania, and the least performing group is Low Income countries. While the values of DA_GOV are lower than D_GOV in each category, the scales differs. Large declines can be observed for Asia, Oceania and the countries that are "not free", and small for Africa and Low Income countries, coinciding with their low D_GOV scores.

5. Discussion

This study contributes to theoretical and practical aspects of digital government benchmarking.

Theoretical considerations lead us to formulate three messages. First, the informal consensus that existing digital benchmarking can contribute to the digital state "metaphor" needs revision. As shown here, this contribution is not obvious. Second, a practice-oriented domain like digital government must rely on contributions from different disciplines. In particular, purely digital conceptualizations are insufficient to capture the impact of digital transformation on the "real" world. Thus digital benchmarking should be applied through an external lens, particularly political one. Third, digital performance is just part of the picture, to be "sustainable", digital benchmarking needs to capture the impact of digital transformation, among other drivers of change, on the "analog" governance [69].

Practical considerations confirm the value of statistical methods for new modes of benchmarking. Problems highlighted in the digital government literature can be quantified and suitable statistical procedures can be applied to address them. It is not necessarily about creating new instruments, even fragmented data coming from existing instruments can be processed to uncover new insights into the outcomes of digitalization. For example, that official propaganda placing strong emphasis on digital advancement can hide serious deficits in analog performance, that digital-analog imbalance shows significant variations between groups of countries, or that impressive results of the national digitalization efforts should be met with some skepticism.

While we expect that this work will benefit theorists and practitioners, the benefits extend to policy-makers and -analysts. This study may help them discover a path to a clear method of evaluating digital strategy performance in relation to intended governance outcomes, and make them aware that relying on assortments of mutually inconsistent instruments may lead to entirely wrong conclusions.

6. Conclusions

This paper demonstrates that existing approaches to digital government benchmarking are not optimal and need a solid revision in regard to their theoretical foundations and practical usage. We have shown that the problems signaled in theoretical studies, such as the diversity and digital-analog disconnect of existing benchmarks, may be expressed in statistical terms, and that established statistical methods exist to help build solutions to them. In particular: 1) specialized methods help reformulate original benchmark information in terms of fewer variables, and 2) it is possible to augment digital measures with additional analog measures for balancing and verification.

This research has some limitations. First, a limited set of indicators was used from two established instruments, spanning one year. A wider selection of instruments and time periods could bring more insights, e.g. uncover evolutionary trends. Second, the statistical methods used come with some limitations that were introduced earlier. Third, a clear interpretation of the constructed synthetic measures may appear challenging. Thus, we treat the results in this paper not as a final product for assessing digital and analog governance, but: 1) as a proof of concept of a method of constructing synthetic and balanced benchmarking instruments, and 2) as elaboration and testing of a logic approach that could drive further efforts towards constructing such instruments.

Fully applicable synthetic and balanced benchmark construction is expected to involve a comprehensive approach including thematic analysis, wide selection of methods, availability of expert opinions, etc. Testing such a construction over time is also necessary. We plan to continue this research to make synthetic and balanced benchmarking directly applicable to digital government theory and practice.

7. References

[1]   F. Fukuyama, "Governance: What Do We Know, and How Do We Know It?," *Annu. Rev. Polit. Sci.*, vol. 19, no. 1, pp. 89–105, 2016.

[2]   A. Bruce-Lockhart, "What do we mean by 'governance'? | World Economic Forum," *World Economic Forum*, 2016. .

[3]   OECD, "OECD Glossary of Statistical Terms - Governance Definition," 2007. .

[4]   T. B. Jørgensen and B. Bozeman, "Public values: An inventory," *Adm. Soc.*, vol. 39, no. 3, pp. 354–381, 2007.

[5]   F. Salem, "Benchmarking the e-government bulldozer: beyond measuring the tread marks," *Meas. Bus. Excell.*, vol. 11, no. 4, pp. 9–22, 2007.

[6]   J. P. Moriarty and C. Smallman, "En route to a theory of benchmarking," *Benchmarking*, vol. 16, no. 4, pp. 484–503, 2009.

[7]   D. Potnis and T. A. Pardo, "Evolution of readiness indicators," in *Proceedings of the 2nd International Conference on Theory and Practice of Electronic Governance, ICEGOV 2008, Cairo, Egypt, December 1-4, 2008*, 2009, p. 417.

[8]   A. Ayanso, D. Chatterjee, and D. I. Cho, "E-Government readiness index: A methodology and analysis," *Gov. Inf. Q.*, vol. 28, no. 4, pp. 522–532, 2011.

[9]   J. Esteves and R. C. Joseph, "A comprehensive framework for the assessment of eGovernment projects," *Gov. Inf. Q.*, vol. 25, no. 1, pp. 118–132, 2008.

[10]   A. Ojo, T. Janowski, and E. Estevez, "Building Theoretical Foundations for Electronic Governance Benchmarking," in *EGOV 2011*, 2011, pp. 13–25.

[11]   E. Siskos, D. Askounis, and J. Psarras, "Multicriteria decision support for global e-government evaluation," *Omega (United Kingdom)*, vol. 46, pp. 51–63, 2014.

[12]   J. Jansen, S. de Vries, and P. van Schaik, "The Contextual Benchmark Method: Benchmarking e-Government services," *Gov. Inf. Q.*, vol. 27, no. 3, pp. 213–219, 2010.

[13]   R. Heeks, "Benchmarking eGovernment: Improving the National and International Measurement, Evaluation and Comparison of eGovernment," 2006.

[14]   D. Tinholt, "e-Government Benchmarks Assessment," 2017.

[15]   F. Bannister, "The curse of the benchmark: An assessment of the validity and value of e-government comparisons," *Int. Rev. Adm. Sci.*, vol. 73, no. 2, pp. 171–188, 2007.

[16]   R. Máchová and M. Lněnička, "Reframing E-government development indices with respect to new trends in ICT," *Rev. Econ. Perspect.*, vol. 15, no. 4, pp. 383–411, 2015.

[17]   L. Al-Hakim, *Global e-government: Theory, applications and benchmarking*. Idea Group Publishing, 2006.

[18]   A. Rorissa, D. Demissie, and T. Pardo, "Benchmarking e-Government: A comparison of frameworks for computing e-Government index and ranking," *Gov. Inf. Q.*, vol. 28, no. 3, pp. 354–362, 2011.

[19]   D. Belanche, L. V. Casalo, and C. Flavian, "The Importance of Confirming Citizens' Expectations in e-Government," in *IFIP Advances in Information and Communication Technology*, 2010, vol. 332, pp. 103–111.

[20]   A. R. M. Schellong, "Benchmarking EU e-government at the crossroads: A framework for e-government benchmark design and improvement," *Transform. Gov. People, Process Policy*, vol. 4, no. 4, pp. 365–385, 2010.

[21]   B. W. Wirtz and P. Daiser, *E-Government: Strategy Process Instruments. Textbook for the Digital Society. 2nd edition*, Speyer. Speyer, 2017.

[22]   G. Hu, W. Pan, M. Lu, and J. Wang, "The widely shared definition of e-Government," *Electron. Libr.*, vol. 27, no. 6, pp. 968–985, 2009.

[23]   OECD, "OECD Glossary of Statistical Terms - E-government Definition," 2002. [Online]. Available: https://stats.oecd.org/glossary/detail.asp?ID=4752. [Accessed: 29-Apr-2019].

[24]   T. Janowski, "Digital government evolution: From transformation to contextualization," *Gov. Inf. Q.*, vol. 32, no. 3, pp. 221–236, 2015.

[25]   D. Janssen, "Mine 's Bigger than Yours : Assessing International eGovernment Benchmarking," in *3rd European Conference on eGovernment*, 2003, pp. 1–12.

[26]   D. Janssen, S. Rotthier, and K. Snijkers, "If You Measure it They will Score: an Assessment of International eGovernment Benchmarking," *Inf. Polity*, vol. 9, no. 3–4, pp. 121–130, 2004.

[27]   E. Ardielli and M. Halásková, "Assessment of E-government in EU countries," *Sci. Pap. Univ. Pardubice*, vol. 22, no. 34, pp. 4–16, 2015.

[28]   N. Cámara and D. Tuesta, "BBVA Research » DiGiX: The Digitization Index," 2017.

[29]   United Nations, "UN Global E-government Survey 2003," 2003.

[30]   OECD, *The e-Government Imperative*. OECD, 2003.

[31]   ITU, "Framework for e-Gov toolkit," pp. 1–35, 2009.

[32]   E. Kromidha, "Strategic e-government development and the role of benchmarking," *Gov. Inf. Q.*, vol. 29, no. 4, pp. 573–581, 2012.

[33]   F. Bannister and R. Connolly, "New Problem for Old? Defining e-Governance," in *Proceedings of the 44th Hawaii International Conference on System Sciences - 2011*, 2011, pp. 1–10.

[34]   J. Freeman, "E-government in the context of monitory democracy: Public participation and democratic reform," *Media Asia*, vol. 40, no. 4, pp. 354–362, 2013.

[35]   D. D. Navarra and T. Cornford, "The State and Democracy After New Public Management: Exploring Alternative Models of E-Governance," *Inf. Soc.*, vol. 28, no. 1, pp. 37–45, 2012.

[36]   V. Bekkers and V. Homburg, "The Myths of E-government: Looking beyond the assumptions of a new and better government," *Inf. Soc.*, vol. 23, no. 5,

pp. 373–382, 2007.

[37] J. Simon, T. Bass, V. Boelman, and G. Mulgan, *NESTA Digital Democracy: the tools transforming political engagement*. 2017.

[38] J. Durkiewicz and T. Janowski, "Is Digitalization Improving Governance Quality? Correlating Analog and Digital Benchmarks," in *ECDG 2018*, 2018, pp. 48–56.

[39] R. M. Gisselquist, "What does good governance mean?," *WIDERAngle*, 2012. .

[40] S. Haque, P. Pathrannarakul, and Boonan, "E-Government towards Good Governance: A Framework for Digital Governance.," *Int. J. Manag. Sci. Bus. Res.*, vol. 3, no. 2, pp. 32–40, 2014.

[41] World Bank, "WGI 2018 Interactive > Home." [Online]. Available: https://info.worldbank.org/governance/wgi/#home. [Accessed: 30-Apr-2019].

[42] N. Saidi and H. Yared, "Fostering Digital Inclusion The Role of ICT in Development eGovernment : Technology for Good Governance , Development and Democracy in the MENA countries," *Mediterr. Dev. Forum IV*, no. April, pp. 1–25, 2002.

[43] V. Madzova, K. Sajnoski, and L. Davcev, "E-Government as an Efficient Tool towards Good Governance," *Balk. Soc. Sci. Rev.*, vol. 1, pp. 157–174, 2013.

[44] S. Suhardi, A. Sofia, and A. Andriyanto, "Evaluating e-Government and Good Governance Correlation," *J. ICT Res. Appl.*, vol. 9, no. 3, pp. 236–262, 2016.

[45] Å. Grönlund, "Connecting eGovernment to real government - The failure of the UN eParticipation index," *Lect. Notes Comput. Sci.*, vol. 6846 LNCS, pp. 26–37, 2011.

[46] Å. Grönlund, "ICT Is Not Participation Is Not Democracy - eParticipation Development Models Revisited," in *ePart 2009*, 2009, pp. 12–23.

[47] G. Lidén, "Technology and democracy: validity in measurements of e-democracy," *Democratization*, vol. 22, no. 4, pp. 698–713, 2015.

[48] S. Stier, "Political determinants of e-government performance revisited: Comparing democracies and autocracies," *Gov. Inf. Q.*, vol. 32, no. 3, pp. 270–278, 2015.

[49] D. K. Basyal, N. Poudyal, and J. W. Seo, "Does E-government reduce corruption? Evidence from a heterogeneous panel data model," *Transform. Gov. People, Process Policy*, vol. 12, no. 2, pp. 134–154, 2018.

[50] S. Baller, S. Dutta, and B. Lanvin, *The Global Information Technology Report 2016 Innovating in the Digital Economy*. 2016.

[51] World Economic Forum, "Technical Notes and Sources." [Online]. Available: http://reports.weforum.org/global-information-technology-report-2016/technical-notes-and-sources/. [Accessed: 07-May-2019].

[52] World Economic Forum, "WEF_NRI_2012-2016_Historical_Dataset." [Online]. Available: http://reports.weforum.org/global-information-technology-report-2016/networked-readiness-index/.

[Accessed: 30-Apr-2019].

[53] C. O. Sorzano, J. Vargas, and A. Pascual-Montano, "A survey of dimensionality reduction techniques based on random projection," no. March. ArXiv, 2014.

[54] I. T. Jolliffe, "Principal Component Analysis. Second Edition," *Springer Ser. Stat.*, vol. 98, p. 487, 2002.

[55] J. F. Hair Jr., W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis*. Harlow: Pearson, 2014.

[56] M. Nardo, M. Saisana, A. Saltelli, and S. Tarantola, *Tools for composite indicators building*. European Comission, 2005.

[57] OECD and JRC, *Handbook of Constructing Composite Indicators: Methodology and user guide*. OECD, 2008.

[58] G. Nicoletti, S. Scarpetta, and O. Boylaud, "Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation," *SSRN Electron. J.*, no. May 2000, 2000.

[59] M. Allen, Ed., "Factor Analysis: Varimax Rotation," in *The SAGE Encyclopedia of Communication Research Methods*, SAGE, 2017.

[60] D. Kaufmann, A. Kraay, and M. Mastruzzi, "The Worldwide Governance Indicators: A Summary of Methodology, Data and Analytical Issues," 2010.

[61] World Bank, "wgidataset." [Online]. Available: https://info.worldbank.org/governance/wgi/#home. [Accessed: 30-Apr-2019].

[62] A. Balicki, *Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne*. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego, 2009.

[63] A. Bąk, "Comparative Analysis Of Selected Linear Ordering Methods Based On Empirical and Simulation Data," *Res. Pap. Wrocław Univ. Econ.*, no. 508, 2018.

[64] C. Distefano, M. Zhu, and D. Mîndrilă, "Understanding and Using Factor Scores: Considerations for the Applied Researcher - Practical Assessment, Research & Evaluation," *Pract. Assessment, Res. Eval.*, vol. 14, no. 20, pp. 1–11, 2009.

[65] Economist Intelligence Unit, "EIU Democracy Index 2018 - World Democracy Report," 2019. [Online]. Available: https://www.eiu.com/topic/democracy-index. [Accessed: 15-Jun-2019].

[66] United Nations, "UNSD — Methodology." .

[67] Freedom House, "Methodology: Freedom in the World 2018 | Freedom House." .

[68] World Bank, "World Bank Country and Lending Groups – World Bank Data Help Desk." .

[69] E. Giovannoni and G. Fabietti, "What Is Sustainability? A Review of the Concept and Its Applications," in *Integrated Reporting: Concepts and Cases that Redefine Corporate Accountability*, no. 2010, C. Busco, M. Frigo, A. Riccaboni, and P. Quattrone, Eds. 2013, pp. 21–40.