

Logical clustering approach for analysing digital economy and society performance of countries

Jovana Rakićević^a and Aleksandar Rakićević^b and Ana Poledica^c

^aDepartment for Technology, Innovation and Development Management, Faculty of Organizational Sciences, University of Belgrade, Jove Ilića 154, jovana.rakicevic@fon.bg.ac.rs

^bDepartment of Systems Theory and Control, Faculty of Organizational Sciences, University of Belgrade, Jove Ilića 154, aleksandar.rakicevic@fon.bg.ac.rs

^cDepartment of Systems Theory and Control, Faculty of Organizational Sciences, University of Belgrade, Jove Ilića 154, ana.poledica@fon.bg.ac.rs

Abstract

Digitization, digitalization and digital transformation represent one of the primary incentives of today's development. To successfully implement these changes, countries need to create smart digital policies which are evidence- and data-based. The study presented in this paper uses the logical clustering approach for grouping countries according to five dimensions of the Digital Economy and Society Index (DESI). Logical clustering employs Interpolative Boolean Algebra (IBA) as a consistent fuzzy approach, which means that all Boolean axioms are fulfilled. To measure proximity among countries, logical clustering uses IBA-based exclusive disjunction and logical aggregation. The general aim is to provide help in identifying directions for defining smart digital policies for achieving digital competitiveness of nations, based on the analysis of similarities among countries. The results indicate that logical clustering enables more comprehensive differentiation between clusters than the original composite index methodology does, and determination of the primary areas of action in clusters, among similar countries. Some interesting cases where logical clustering results differ from the original methodology are discussed.

Keywords: Logical clustering, Interpolative Boolean Algebra, Consistent fuzzy logic, Digital Economy and Society Index, Digital transformation, Digital policy.

1 Introduction

Industrial revolutions which occurred in the last three centuries caused a profound change in economic

systems and social structures and shaped novel ways of living. Today we are witnessing the powerful fourth industrial revolution which is led by artificial intelligence, machine learning and the Internet of things (IoT) [26]. This era of digital transformation puts new challenges in front of governments, companies, entrepreneurs, and customers. Switching from health to e-health, from government to e-government, from traditional to digital business models, from cities to smart cities, became some of the key activities in modern societies. This rapid digital "e"-wave imposes the need of directing governments towards the creation of smart digital policy mix that will enable digitization, digitalization, and finally the digital transformation of nations through a set of balanced measures and instruments. To properly manage this process, measuring and monitoring the digital performance of countries emerged as "a must".

In this regard, both international organizations and individual academic researchers are trying to develop measures of digital performance of countries [2, 3, 4, 12, 13]. Most of these measures are based on the composite index methodology [16]. Numerous composite indices have evolved lately in various areas of nations' development and became a very popular way of assessing large phenomena that cannot be captured by one single indicator. They serve as a tool for measuring countries performance and ranking accordingly, enabling the cross-country comparisons and facilitation of the strategic decision-making processes [11]. Analysing results of these measures lead to better, fine-tuned policies and strategies at various levels of the economy [24]. Examples include assessments of human development, innovativeness, competitiveness, sustainability, wellbeing, and many others (for a review see [1]). Recent digital trends caused another index to emerge – the Digital Economy and Society Index (DESI), developed by the European Commission in 2014 for tracking the digital performance of the EU countries. This index is in the centre of the analysis presented in this paper.

However, although composite indices overwhelmed

the field of performance measurement of countries, there are many obvious and often criticised shortcomings. First and the most controversial step in creating composite indices is the methodology used for weighting and aggregating values [25], which is most usually simple or weighted average function [8]. Secondly, it has become increasingly difficult to select appropriate individual indicators feeding into the final index due to easily accessible information overload. Furthermore, the collected data on countries' performance are often imprecise, missing, and are both quantitative and qualitative in nature. This supports the identified need to extending traditional decision-making processes by including intuitive reasoning and modelling of uncertainty, vagueness, and imprecision for properly facilitating the decision-making process at various levels.

This paper addresses this issue by using the logical clustering approach for analysing the digital performance of 28 EU member countries with the aim to open a path towards the development of smart digital policy mix of countries. The clustering problem has been addressed in many contexts and by researchers in many disciplines, indicating its usefulness as one of the steps in exploratory data analysis. Clustering countries according to performance showed valuable results in identifying common patterns of countries' development in many areas (see e.g. [6, 7, 14]). Logical clustering technique used in this paper, which employs Interpolative Boolean Algebra as a consistent fuzzy approach, enables more comprehensive differentiation between clusters [21].

This paper has two main research directions. Firstly, we understand that clustering countries according to digital performance can serve as a help in directing governments towards identifying smart data-driven and evidence-based digital policy mix. Secondly, from the methodological viewpoint, we try to detect to which extent logical clustering approach which includes intuitive reasoning differs from the original composite index DESI methodology which is based on a simple weighted function. Some interesting cases where differences are noticed are highlighted.

The paper is organized as follows. Section 2 introduces DESI which is used for measuring the digital performance of countries. Section 3 explains the logical clustering approach. Section 4 presents the application of logical clustering to DESI index (data collection and preparation, and results and discussion). Section 5 concludes.

2 Digital Economy and Society Index

DESI measures the progress of 28 EU member states towards digital economy and society. It tracks the evolution of countries in digital competitiveness, bringing together a set of relevant indicators on Europe's current digital policy mix [4].

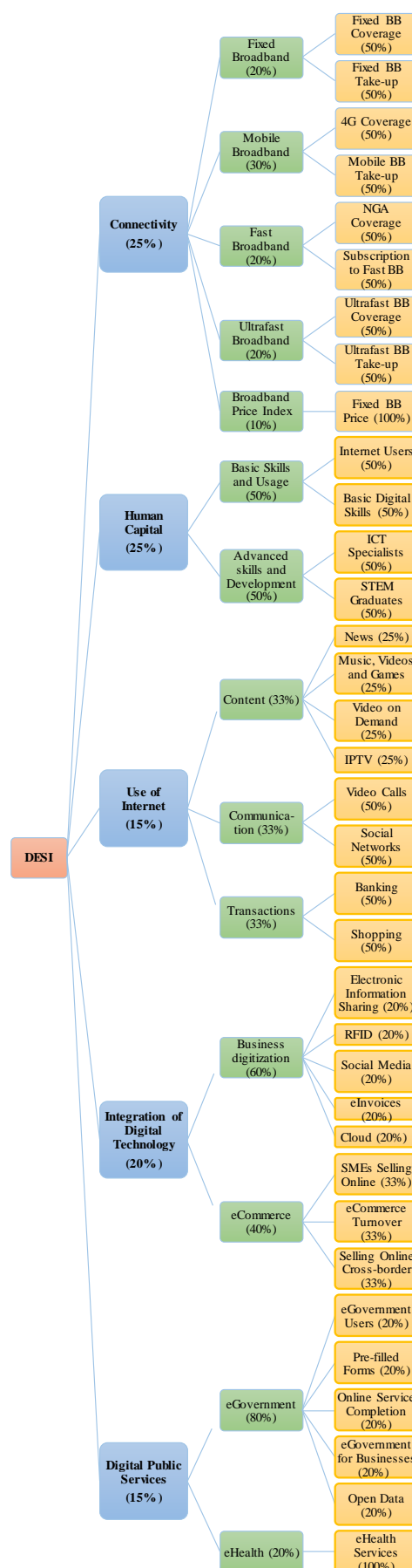


Figure 1: The DESI structure (based on [4])

DESI is composed of five main dimensions: Connectivity, Human Capital, Use of Internet, Integration of Digital Technology, and Digital Public Services. Each dimension is divided into a set of sub-dimensions, which are in turn composed of individual indicators [4]. Since its first introduction in 2014, the number of individual indicators covered by the final DESI index varied. From 30 indicators aggregated for obtaining the final index value in 2014, the list was expanded to 35 indicators in 2018.

The aggregation of indicators into sub-dimensions, of sub-dimensions into dimensions, and of dimensions into the overall DESI index is performed bottom-up using simple weighted arithmetic averages following the structure of the index. The three-layer index structure with the defined weights of elements is presented in Figure 1.

3 Research method: Logical clustering

Here we explain logical clustering (LC), a consistent fuzzy approach to clustering method [21, 22].

3.1 Theoretical background

LC is theoretically based on interpolative Boolean Algebra (IBA) [17] as being a Boolean-consistent fuzzy approach [18]. More precisely, it means that all Boolean laws are fulfilled in $[0,1]$ interval including the laws of contradiction and excluded middle.

Another IBA-based method that is important for LC approach is logical aggregation (LA) [19] that enables including logical dependencies between attributes in the process of aggregation. LA is a very powerful aggregation tool since it generalizes many conventional aggregation operators such as weighted sum, arithmetic mean, OWA, etc. [19]. From the aspect of application, one of the main benefits of LA is its interpretability, since LA functions are easy to understand and analyse [21]. Further, the previous benefits are integrated and serve as a basis for providing an LC approach.

3.2 Logical clustering method

Clustering, defined as the “unsupervised classification of patterns (observations, data items, or feature vectors) into groups (clusters)” [9], is a method of extracting knowledge about the hidden structures in the data, partitioning it into optimally homogeneous groups on the basis of empirical measures of similarity among those objects [10]. The core of each clustering algorithm is the proximity measure or method used for object comparison. However, no universal measure can consider different aspects of proximity. Instead, an appropriate one from the wide set of different measures of distance, similarity/dissimilarity or correlation must be chosen (for a review on different measures of proximity see [5]).

Logical clustering, firstly introduced in [22], is an IBA-based clustering technique within a standard hierarchical clustering algorithm. As being a consistent fuzzy approach, LC enables more comprehensive differentiation between clusters [21]. Some of the applications of LC approach is found in [21, 22, 23].

LC algorithm consists of the following steps:

- (1) **Data normalization** on $[0,1]$ interval is a mathematical prerequisite for processing data in IBA which serves as a basis for LC. Standard min-max normalization is often used. Depending on the nature of the problem/attributes more specific normalization functions may be interesting for application, especially the ones that change not only range but also data distribution.
- (2) The second step refers to the **object comparison** using IBA relations. Hierarchical clustering groups objects organizing them into a hierarchical structure according to their proximity. One of the key parts of the LC approach is a logical measure of proximity i.e. IBA relation of exclusive disjunction [17, 22]:

$$\begin{aligned}(a \vee b)^{\otimes} &= ((a \wedge \neg b) \vee (\neg a \wedge b))^{\otimes} = \\ &= a + b - 2 \cdot a \otimes b = \\ &= a + b - 2 \cdot \min(a, b)\end{aligned}$$

where a and b are the attributes of two objects A and B . Considering that a and b are the same attributes (that is of similar nature), a generalized product \otimes is replaced with minimum function [21].

As being a complementary relation of an equivalence relation, IBA exclusive disjunction is an intuitive measure of similarity/dissimilarity which is close to human perception [20, 21]. It is used to calculate proximity between object attributes. To obtain overall proximity between objects as a single representative value, we apply various logical aggregation operators. The following is the disjunction function that is used as a proximity measure in this paper:

$$d_{IBA}^{disj}(A, B) = (a_1 \vee b_1) \vee (a_2 \vee b_2) \vee \dots \vee (a_n \vee b_n)$$

where $A = [a_1, \dots, a_n]$ and $B = [b_1, \dots, b_n]$ are two multi-attribute objects.

In [15] it is shown that the IBA framework for modelling similarity/dissimilarity covers a wide range of similarity/dissimilarity measures, making it suitable for various problems.

- (3) Referring to the **application of a linkage criterion**, the LC algorithm uses the group-average linkage method.

- (4) **The number of clusters** is typically determined by expert judgement using dendrogram.

4 Data and results

In the study, we employ LC based on agglomerative hierarchical approach over five dimensions of DESI: Connectivity, Human Capital, Use of Internet, Integration of Digital Technology, and Digital Public Services, for clustering countries based on their digital performance.

4.1 Data collection and preparation

The study is performed on a sample of 28 EU member countries over the period 2017-2018. Data is collected from the European Commission website¹. Here we emphasize that raw data used in this study refers to the values of five DESI dimensions which are obtained by aggregating values of individual indicators following DESI structure presented in Figure 1. To exemplify, value for the Connectivity is obtained by aggregating values of sub-dimensions (Fixed Broadband, Mobile Broadband, Fast Broadband, Ultrafast Broadband, Broadband Price Index), which actually represent aggregated values of individual indicators. As an illustration, only a part of the raw data for 2017 is presented in Table 1.

Country	Connectivity	Human Capital	Use of Internet	Integration of Digital Technology	Digital Public Services
1 Austria	58,81	62,44	43,89	39,35	66,28
2 Belgium	72,70	57,31	51,85	52,39	52,26
3 Bulgaria	51,55	31,09	38,60	22,49	45,18
...
27 Sweden	72,48	69,30	71,42	53,83	67,45
28 UK	64,04	71,34	59,43	36,95	56,18

Table 1: Raw data for 2017

Since LC uses the logical measure of similarity/dissimilarity, we need to normalize the raw data to [0,1] interval. In this study we choose the standard min-max normalization function:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Normalized data that corresponds to the raw data from Table 1 is presented in Table 2.

Country	Connectivity	Human Capital	Use of Internet	Integration of Digital Technology	Digital Public Services
1 Austria	0.49974	0.68924	0.33119	0.474	0.74528
2 Belgium	0.86426	0.57703	0.50878	0.77195	0.42553
3 Bulgaria	0.30926	0.00461	0.21326	0.08850	0.26371
...
27 Sweden	0.85852	0.839	0.94518	0.80485	0.77236
28 UK	0.63698	0.88359	0.67771	0.41897	0.5151

Table 2: Normalized data for 2017

4.2 Results and discussion

Tables 3 and 6 show the results of LC for 2018 and 2017, respectively. In both tables, the structure is as follows. Column 1 presents the analysed countries. Column 3 shows the final DESI values, while column 2 presents the difference in those values between two subsequent countries. Values marked in bold indicate the largest gaps. Column 4 ranks the countries according to the values presented in column 3. Columns 5 and 6 show the results of LC of countries according to the digital economy and society performance. More precisely, column 5 presents the results of LC with three clusters (LC3), while column 6 presents LC results with five clusters and two outliers (LC5+2).

2018		DESI		Rank	LC3	LC5+2
(1)		(2)	(3)	(4)	(5)	(6)
Denmark	DK		71.73	1	1	1
Sweden	SE	1.28	70.45	2	1	1
Finland	FI	0.34	70.11	3	1	1
Netherlands	NL	0.24	69.87	4	1	1
Luxembourg	LU	7.08	62.79	5	2	2
Ireland	IE	1.52	61.26	6	2	3
UK	UK	0.06	61.21	7	2	2
Belgium	BE	0.47	60.73	8	2	3
Estonia	EE	0.99	59.74	9	2	*
Spain	ES	1.69	58.05	10	2	3
Austria	AT	0.09	57.96	11	2	3
Malta	MT	0.30	57.66	12	2	2
Lithuania	LT	1.05	56.61	13	2	3
Germany	DE	0.99	55.61	14	2	4
Slovenia	SI	2.61	53.00	15	2	4
Portugal	PT	0.42	52.59	16	2	4
Czech Rep.	CZ	0.26	52.32	17	2	4
France	FR	0.80	51.53	18	2	4
Latvia	LV	0.68	50.84	19	2	**
Slovakia	SK	1.37	49.48	20	2	4
Cyprus	CY	0.14	49.34	21	2	4
Croatia	HR	2.68	46.66	22	2	4
Hungary	HU	0.12	46.55	23	2	4
Poland	PL	1.52	45.02	24	3	5
Italy	IT	0.77	44.25	25	3	5
Bulgaria	BG	3.22	41.03	26	3	5
Greece	EL	2.65	38.38	27	3	5
Romania	RO	0.83	37.55	28	3	5

Table 3: Results for 2018

As far as 2018 is concerned, we firstly used the difference in the original DESI values (see column 2) to identify if considerable gaps exist between any two subsequent countries. In other words, to identify the last country of one group and the first of the subsequent. Based on that information, we identified three groups of countries: leaders (from 1 to 4 ranking); average (from 5 to 25); falling behind (from 26 to 28). These three clusters are simply identified by analysing column 2 where bolded values indicate the largest gaps in DESI values. This step was performed to examine if we could determine the clusters of countries by simply analysing the

¹ All the data is publicly available online at: <https://digital-agenda-data.eu/datasets/desi/visualizations>

composite index results. Understanding that three clusters emerge from these results, we performed LC approach with three clusters, to see whether the results of LC differ and to what extent.

Results of LC with three groups (see column 5) slightly differ from the original results. The first cluster of most developed countries remains the same. The difference is noticed with Poland and Italy. According to the original results, those two countries are a part of cluster 2. According to the LC approach, they are grouped together with Bulgaria, Greece and Romania in the third cluster. Thus, further analysis of these two countries' scores is performed to understand the difference in results. We firstly calculated the average values of clusters 2 and 3 without Poland and Italy, marked as C2' (19 countries) and C3' (3 countries) respectively. Then, we calculated the absolute differences between these two countries' values and average values of clusters C2' and C3', by each dimension of DESI. These results are shown in Table 4, where column 7 reflects the sum of absolute differences between the cluster's average and the countries' values by all five dimensions of DESI. According to these results, both Poland and Italy are closer to cluster 3, since the difference in digital performance is higher when compared to the cluster 2. Namely, for Poland difference of 1.242 is higher than 0.628, thus Poland is closer to the third cluster. Italy performs similar ($1.261 > 0.771$). The situation remains the same when the weights of DESI dimensions are included (see column 8). The differences for each dimension are multiplied by the weights of those dimensions and then summed into the weighted difference value (see the weights of five DESI dimensions in Figure 1). Weighted difference is higher when comparing these two countries with the average of cluster 2. Thus, LC results, where Poland and Italy are grouped with Bulgaria, Greece and Romania in the third cluster are favoured over the original results in terms of similarity between countries.

	Conn.	Hum. Cap.	Use of Int.	Int. of Dig. Tech.	Dig. Pub. Serv.	Diff.	W-Diff.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PL-C2'	0.149	0.152	0.276	0.399	0.265	1.242	0.236
PL-C3'	0.178	0.281	0.036	0.011	0.121	0.628	0.141
IT-C2'	0.307	0.311	0.394	0.094	0.156	1.261	0.256
IT-C3'	0.021	0.123	0.081	0.317	0.230	0.771	0.146

Table 4: Poland (PL) and Italy (IT) analysis (case of LC3), year 2018

Furthermore, from Table 3 we notice that the second cluster of countries is too large. Concrete conclusions cannot be derived, nor the fine-tuning of the digital policies according to cluster affiliation. Thus, we perform the LC with five clusters to differentiate better among these countries. Estonia and Latvia are identified as outliers. First analysis shows that these two countries deviate the most from other countries in the corresponding clusters, analysing the pairwise deviations between countries. By changing the cutting parameter, these two countries would become

members of clusters 3 and 4 respectively, which calls for further analysis. However, results show a few interesting cases with clusters 2 and 3 in LC5+2 results (see column 6, Table 3). From the column 6 of Table 3 we see that these two clusters intertwine. Namely, Malta which is ranked 12th according to DESI belongs to the second cluster, together with Luxembourg (ranked 5th) and the United Kingdom (ranked 7th). Malta's score is further examined in a similar way to Poland and Italy in the previous case. Results from Table 5 indicate that Malta is closer to the average of the second cluster, both by the difference ($0.632 < 0.974$) and the weighted difference ($0.143 < 0.189$). It is interesting to notice that Malta is closer to the average of the second cluster by all DESI dimension except the Human Capital (see columns 2-6 Table 5). Connectivity, Use of Internet and Integration of Digital Technology are almost at the same level as the average of cluster 2. Thus, by analysing the similarity/dissimilarity between countries, Malta should be a part of the second cluster. In the same manner, Ireland which is ranked 6th is actually a part of the third cluster according to the LC approach (see column 6, Table 3).

	Conn.	Hum. Cap.	Use of Int.	Int. of Dig. Tech.	Dig. Pub. Serv.	Diff.	W-Diff.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MT-C2'	0.034	0.421	0.021	0.054	0.102	0.632	0.143
MT-C3'	0.169	0.121	0.284	0.282	0.118	0.974	0.189

Table 5: Malta analysis (case of LC5+2), year 2018

Even though these countries are in the same cluster according to original DESI methodology (see column 5, Table 3), the dendrogram (Figure 2) clearly illustrates that these countries are distant from each other and there is no possibility that they can end up in the same cluster.

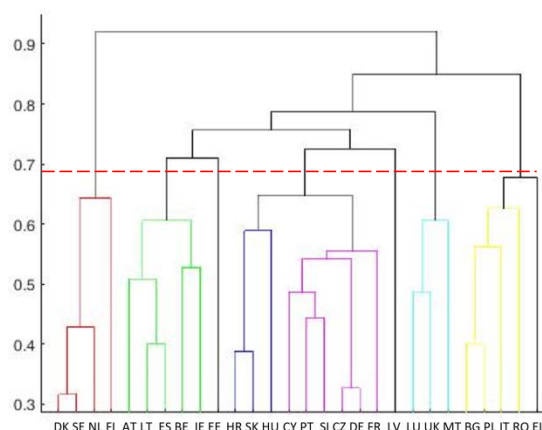


Figure 2: Dendrogram for 2018 (case of LC5+2)

This is an interesting spot in the analysis if we understand the general purpose of the paper. For example, if the European Union decides to direct an action or a policy instrument or an investment for boosting digital competitiveness in the top 10 DESI countries, Malta would be excluded. The LC, however, shows that Malta (ranked 12th by DESI

results) is similar even to Luxembourg which is ranked 5th. Thus, deciding to invest in certain clusters, rather than simple ranking by composite index methodology, would provide in-depth results.

A similar analysis can be performed for 2017 (Table 6). Compared with the analysis for 2018 (Table 3), it is interesting that LC with three clusters (see column 5, Table 6) shows different results from the original methodology (see column 3, Table 6) in 2017.

2017		DESI	Rank	LC3	LC5+2
(1)		(2)	(3)	(4)	(5)
Denmark	DK	70,13	1	1	1
Finland	FI	2,93	67,20	2	1
Sweden	SE	0,16	67,04	3	1
Netherlands	NL	0,50	66,54	4	1
Luxembourg	LU	6,17	60,37	5	1
Belgium	BE	1,77	58,60	6	2
UK	UK	0,02	58,58	7	1
Estonia	EE	1,61	56,96	8	2
Ireland	IE	0,62	56,34	9	2
Malta	MT	1,41	54,93	10	2
Austria	AT	0,22	54,71	11	2
Spain	ES	1,37	53,34	12	2
Lithuania	LT	0,14	53,20	13	2
Germany	DE	0,51	52,69	14	2
Portugal	PT	2,00	50,69	15	2
Slovenia	SI	0,31	50,38	16	2
Czech Rep.	CZ	1,04	49,35	17	2
France	FR	0,50	48,84	18	2
Latvia	LV	1,62	47,22	19	3
Slovakia	SK	1,76	45,46	20	3
Cyprus	CY	0,24	45,22	21	3
Hungary	HU	0,99	44,24	22	3
Croatia	HR	1,05	43,19	23	3
Poland	PL	1,10	42,08	24	3
Italy	IT	0,64	41,44	25	3
Bulgaria	BG	3,71	37,73	26	3
Greece	EL	2,18	35,54	27	3
Romania	RO	1,81	33,73	28	3

Table 6: Results for 2017

Following LC3 approach, Luxembourg (ranked 5th) and the United Kingdom (ranked 7th) are grouped into the first cluster of most developed countries, together with Denmark, Finland, Sweden, and the Netherlands. The DESI report identifies first four countries as a group of most developed. If the European Union decides to direct the funds in the cluster of most developed countries, Luxembourg and the United Kingdom would be omitted following the original methodology. And should they? Table 7 shows that Luxembourg (ranked 5th) is more similar to the United Kingdom (ranked 7th) than to Belgium (6th) (see columns 2 and 3). Both the difference and weighted difference is higher between Luxembourg and Belgium than between Luxembourg and the United Kingdom ($1.387 > 0.875$ and $0.282 > 0.180$). Furthermore, if we calculate the average value of cluster 1' (consisted of Denmark, Finland, Sweden,

Netherlands, and Luxembourg), marked as C1', we can determine which of these two countries is more similar to the first cluster. From columns 5 and 6 (Table 7) we can notice that the United Kingdom is more similar to the average of this cluster than Belgium is, even though according to the original DESI methodology, Belgium is ranked better, and thus is "closer" to the first cluster than the United Kingdom is following the original DESI (column 4, Table 6). Dendrogram presented in Figure 3 graphically supports these conclusions.

	Diff.	W-Diff.		Diff.	W-Diff.
(1)	(2)	(3)	(4)	(5)	(6)
LU-BE	1.387	0.282	BE-C1'	1.039	0.193
LU-UK	0.875	0.180	UK-C1'	0.942	0.183

Table 7: Luxembourg (LU), Belgium (BE) and the UK analysis (case of LC3), year 2017

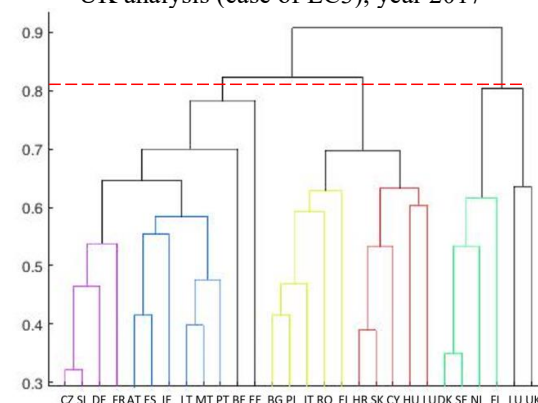


Figure 3: Dendrogram for 2017 (case of LC3)

Looking at 2017, we performed the LC5+2 approach also (column 6, Table 6), to differentiate among the countries in the second cluster. These results support the previous results of LC5+2 for 2018, and will not be examined in detail.

Furthermore, cluster analysis enables us to identify the priority areas of fast action in which countries of the same cluster differ most. This can serve as a help for those less developed in a cluster to transfer good practices from the more developed but similar countries, supporting the evidence-based policy making. Table 8 presents the results for both 2018 and 2017 for LC5+2 approach with five clusters and 2 outliers.

Cluster	No. elem.	Conn.	Hum. Cap.	Use of Int.	Int. of Dig. Tech.	Dig. Pub. Serv.
2018						
1	4	0.173	0.077	0.121	0.097	0.096
2	3	0.149	0.243	0.045	0.084	0.065
3	5	0.124	0.131	0.089	0.143	0.135
4	9	0.145	0.133	0.107	0.141	0.165
5	5	0.166	0.132	0.102	0.160	0.144
2017						
1	4	0.142	0.078	0.139	0.136	0.092
2	2	0.257	0.029	0.070	0.115	0.149
3	10	0.103	0.139	0.135	0.130	0.182
4	5	0.177	0.112	0.044	0.130	0.178
5	5	0.132	0.136	0.113	0.124	0.140

Table 8: Identifying priority areas of difference in clusters (case of LC5+2)

The elements in the table represent the standard deviation of differences between values of countries of a cluster with the average of an affiliated cluster. Values in bold indicate the areas of largest divergence by dimensions of DESI. To exemplify, for 2018 we see that regarding cluster 1, Connectivity and Use of Internet are the primary areas of action with standard deviations of 0.173 and 0.121 respectively. Looking more deeply into the original normalized data (a part of the data is presented in Table 2), we see that in the first cluster the Netherlands is the most developed country regarding Connectivity (1.0) and Finland is the less developed (0.60347). Thus, Finland could transfer good practices from the Netherlands as the primary steps for enhancing digital competitiveness. In the second cluster in 2018 (consisted of Luxembourg, Malta and the United Kingdom), Human Capital imposes as the priority area. Looking into the original data, Malta should boost the number of Internet users, Basic digital skills, ICT specialists, and STEM graduates (elements of the Human Capital pillar) since it is placed at the bottom of the cluster regarding this dimension in comparison to the UK and Luxembourg which lead this cluster. In this manner, all the clusters and all the dimensions can be analysed, and concrete digital policies, measures and instruments can be cascaded down to the individual indicators' improvement of each country. Thus, LC analysis which is based on identifying groups of "similar" countries proves to be useful in identifying smart data- and evidence-based digital policies.

6 Conclusion

The study presented in this paper uses a consistent fuzzy approach to clustering which is based on IBA theory, i.e. the logical clustering (LC) approach. LC is used for providing help in the first steps towards the creation of smart digital policy mix of countries for achieving higher digital competitiveness. The EU countries are clustered according to five main dimensions of DESI which is the most comprehensive index of digital performance of countries. DESI addresses the five principal policy areas of concern for a digital economy and society. Those are Connectivity, Human Capital, Use of Internet, Integration of Digital Technology, and Digital Public Services.

The results indicate the following. Firstly, the LC approach provides different results to the original composite index methodology (DESI) which is typically based on the weighed sum. By respecting the similarity/dissimilarity of countries through the logical measure of proximity together with an operator of logical aggregation, LC includes intuitive reasoning into the clustering process. By using the LC approach, it is possible to model interaction among DESI components but also the effect of compensation among the index counterparts. Secondly, due to these reasons, LC proves to be

useful in identifying more comprehensive clusters of countries. This further enables the fine-tuning of digital policies, measures, and instruments to a group of similar countries in terms of digital competitiveness – leaders, average, and falling behind countries. Thirdly, it is possible to determine first areas of fast action in each cluster of similar countries, taking into account the areas of largest divergence by dimensions of DESI. Finally, by determining leaders and latecomers in each cluster of similar countries, respecting all dimensions of the DESI, it is possible to provide a basis for replicating and adjusting policies, measures and instruments and transferring good practices from more developed countries, but similar once in terms of cluster affiliation, to those less developed in a cluster. Concrete cases which confirm these statements are presented in the results and discussion section.

From the aspect of LC method, future analysis can be directed towards finding the more appropriate operator of logical aggregation for the analysed problem. Bearing in mind the composite multiple-layer structure of DESI, we could apply more sophisticated techniques of aggregation (e.g. logical aggregation) on all levels of the DESI index structure in future research.

References

- [1] R. Bandura, Composite Indicators and Rankings: Inventory 2011. New York: United Nations Development Programme, Office of Development Studies (UNDP/ODS Working Paper), 2011.
- [2] M. Billon, F. Lera-Lopez, R. Marco, Differences in digitalization levels: a multivariate analysis studying the global digital divide, *Review of World Economics* 146(1) (2010) 39-73.
- [3] N. Corrocher, A. Ordanini, Measuring the digital divide: A framework for the analysis of cross-country differences, *Journal of Information technology* 17(1) (2002) 9-19.
- [4] European Commission, DESI 2018 - Digital Economy and Society Index Methodological note, Directorate-General for Communications Networks, Content and Technology, 2018.
- [5] B.S. Everitt, S. Landau, M. Leese, D. Stahl, *Cluster Analysis* 5th ed, John Wiley & Sons Ltd., Chichester, 2011.
- [6] A. Filippetti, A. Peyrache, The patterns of technological capabilities of countries: a dual approach using composite indicators and data envelopment analysis, *World Development* 39(7) (2011) 1108-1121.

- [7] F. Franceschini, M. Galetto, D. Maisano, L. Mastrogiacomo, Clustering of European countries based on ISO 9000 certification diffusion, *International Journal of Quality & Reliability Management* 27(5) (2010) 558-575.
- [8] H. Grupp, T. Schubert, Review and new evidence on composite innovation indicators for evaluating national performance, *Research Policy* 39 (2010) 67-78.
- [9] A. K. Jain, M. N. Murty, P. J. Flynn, Data clustering: a review, *ACM computing surveys*, 31(3) (1999) 264-323.
- [10] S. C. Johnson, Hierarchical clustering schemes, *Psychometrika* 32(3) (1967), 241-254.
- [11] M. Jovanović, J. Rakićević, M. Levi Jakšić, J. Petković, S. Marinković, Composite Indices in Technology Management: A Critical Approach, in: V. Jeremić, Z. Radojčić, M. Dobrota (Eds.), *Emerging Trends in the Development and Application of Composite Indicators*, IGI Global, Hershey, 2017, pp. 38-71.
- [12] R. Katz, P. Koutroumpis, Measuring digitization: A growth and welfare multiplier, *Technovation* 33 (10-11) (2013) 314-319.
- [13] R. Katz, P. Koutroumpis, F. Martin Callorda, Using a digitization index to measure the economic and social impact of digital agendas, *Info* 16(1) (2014) 32-44.
- [14] S. Kaufmann, The business cycle of European countries Bayesian clustering of country-individual IP growth series (No. 83), *Austrian Central Bank* 2003.
- [15] P. Milosevic, A. Poledica, A. Rakicevic, V. Dobric, B. Petrovic, D. Radojevic, IBA-based framework for modeling similarity, *International Journal of Computational Intelligence Systems* 11 (2018) 206-218.
- [16] M. Nardo, M. Saisana, A. Saltelli, S. Tarantola, A. Hoffman, E. Giovanni, *Handbook on constructing composite indicators: methodology and user guide*, OECD publishing, Paris, 2008.
- [17] D. Radojevic, New [0,1]-valued logic: A natural generalization of Boolean logic, *Yugoslav Journal of Operational Research* 10(2) (2000) 185-216.
- [18] D. Radojevic, Interpolative Realization of Boolean Algebra as a Consistent Frame for Gradation and/or Fuzziness, in: M. Nikraves, J. Kacprzyk, L. Zadeh (Eds.), *Forging New Frontiers: Fuzzy Pioneers II*, Springer-Verlag Berlin Heidelberg, 2008a, pp. 295-318.
- [19] D. Radojevic, Logical aggregation based on interpolative Boolean algebra, *Mathware & Soft Computing* 15 (2008b) 125-141.
- [20] A. Rakićević, V. Dobrić, D. Radojević, Selection of equity securities with logical aggregation, in: D. Ruan, L. Tianrui, C. Guoqing, X. Yang (Eds.), *Foundations and Applications of Computational Intelligence. Proc. 9th International FLINS Conference*, Chengdu, 2010, pp. 603-609.
- [21] A. Rakićević, P. Milošević, A. Poledica, I. Dragović, B. Petrović, Interpolative Boolean Approach for Fuzzy Portfolio Selection, in: A. Meier, E. Portmann, L. Terán (Eds.), *Applying Fuzzy Logic for the Digital Economy and Society, Fuzzy Management Methods*, Springer, Cham, 2019, pp. 23-46.
- [22] A. Rakićević, I. Nešić, D. Radojević, A novel approach to hierarchical clustering based on logical measure of dissimilarity, in: N. Mladenović, G. Savić, M. Kuzmanović, D. Makajić-Nikolić, M. Stanojević (Eds.), *Proceedings of the 11th Balkan Conference on Operational Research (BALCOR 2013)*, UB FOS, Zlatibor, 2013, pp. 147-155.
- [23] A. Rakićević, V. Simeunović, B. Petrović, S. Milić, An Automated System for Stock Market Trading Based on Logical Clustering, *Tehnički vjesnik* 25(4) (2018) 970-978.
- [24] J. Rakićević, M. Levi Jakšić, M. Jovanović, Measuring the Potential for Technology Entrepreneurship Development: Serbian Case, *Management: Journal of Sustainable Business and Management Solutions in Emerging Economies* 23(2) (2018) 13-25.
- [25] M. Saisana, A. Saltelli, S. Tarantola, Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society*, 168 (2) (2005) 307-323.
- [26] K. Schwab, *The Fourth Industrial Revolution*, World Economic Forum, Geneva, 2016