

A “posetic” toolbox for the construction of synthetic indicators

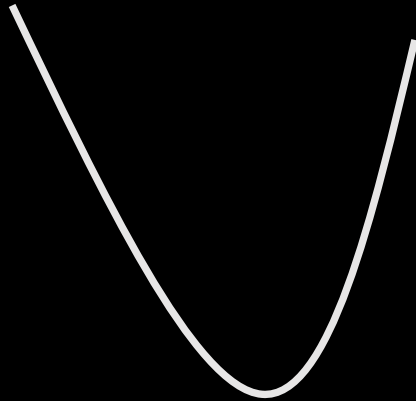
Marco Fattore, Filomena Maggino

JRC – March 23th, 2018

1. Yesterday we talked about the nature, the use and the communication of indicators.
2. Among other things, we said indicators should preserve complexity and should provide a faithful representation of reality.
3. Now, going to the “back-office” (the “factory”):
 - How these concepts relate to the way we develop indicators? What principles drive us in this process?
 - Are we, in the end, capable to provide “dense information and complexity preserving indicators”, for a complex world?
 - Which are (some of) the conceptual and technical challenges we must face, in order not to be “data analyst”, but “**information condenser**”?
 - What are we working on, in practice?

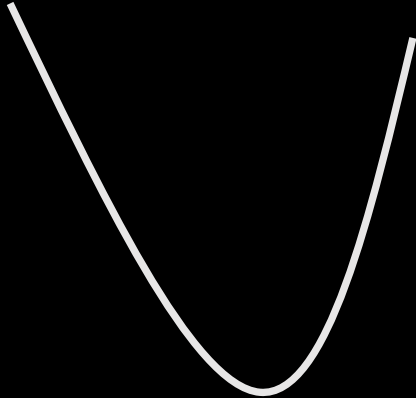
Problem: prove that the minimum is unique

GEOMETRIC WORLD

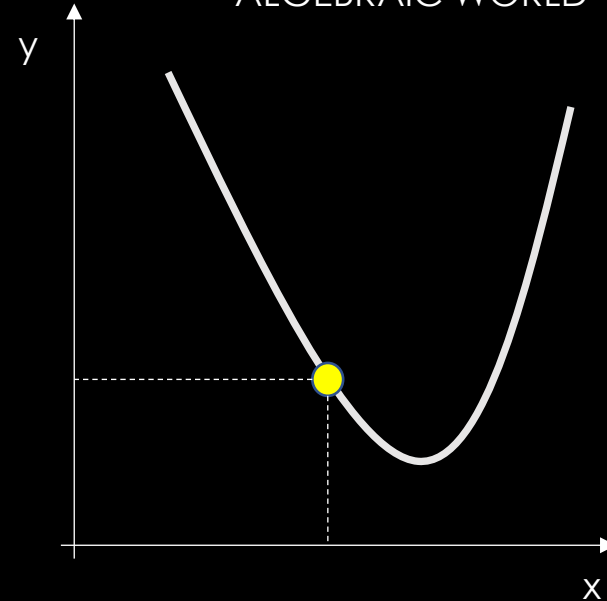


Problem: prove that the minimum is unique

GEOMETRIC WORLD



ALGEBRAIC WORLD



ALGEBRAIC **REPRESENTATION**

GEOMETRIC TRUTH



$y = f(x)$
put $f'(x) = 0$ and show that the solution is unique



ALGEBRAIC TRUTH

Structure preserving maps

1. Map «parts of reality» into the proper data structure (**representation**)

Assumption: truths about reality are «loaded» implicitly into the data structure.

2. Elaborate on the data structure using tools and procedure which are **consistent** with it (the right «syntax»)

Assumption; this makes explicit «hidden truths» about the data.

3. Get **back to reality**, by mapping back from the outputs (semantics).

Assumption: the extracted “hidden truths” have **something to say** about reality (not more than something).

AIRBAG Life Cycle Assessment

The problem of partially ordered data

There are some «holes» in the statistical literature, so that currently we do not know how to treat some kinds of data structures, mainly

DATA WHICH ARE TRULY MULTIDIMENSIONAL AND CANNOT BE
COMPLETELY ORDERED

i.e.

PARTIALLY ORDERED DATA

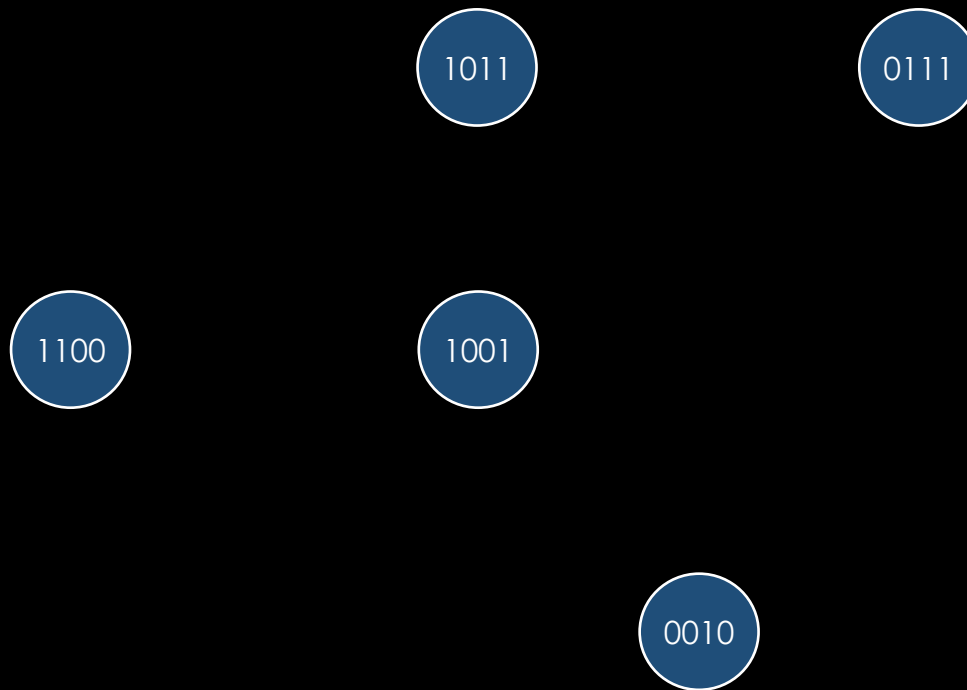
Partially ordered data are spreading, often, but not exclusively, under the form of Multidimensional Systems of Ordinal Attributes., particularly in the socio-economic field.

Child wellbeing in Congo (4 binary variables)

- Sanitation deprivation Access to any kind of improved latrines or toilets
- Water deprivation Only access to surface water for drinking or the nearest source of water is more than a 15 minutes walking distance from the dwellings
- Shelter deprivation Living in dwellings with more than five people per room or with no flooring material (e.g.. a mud floor)
- Health deprivation The nearest health service provider is more than a 15 minutes walking distance from their dwellings

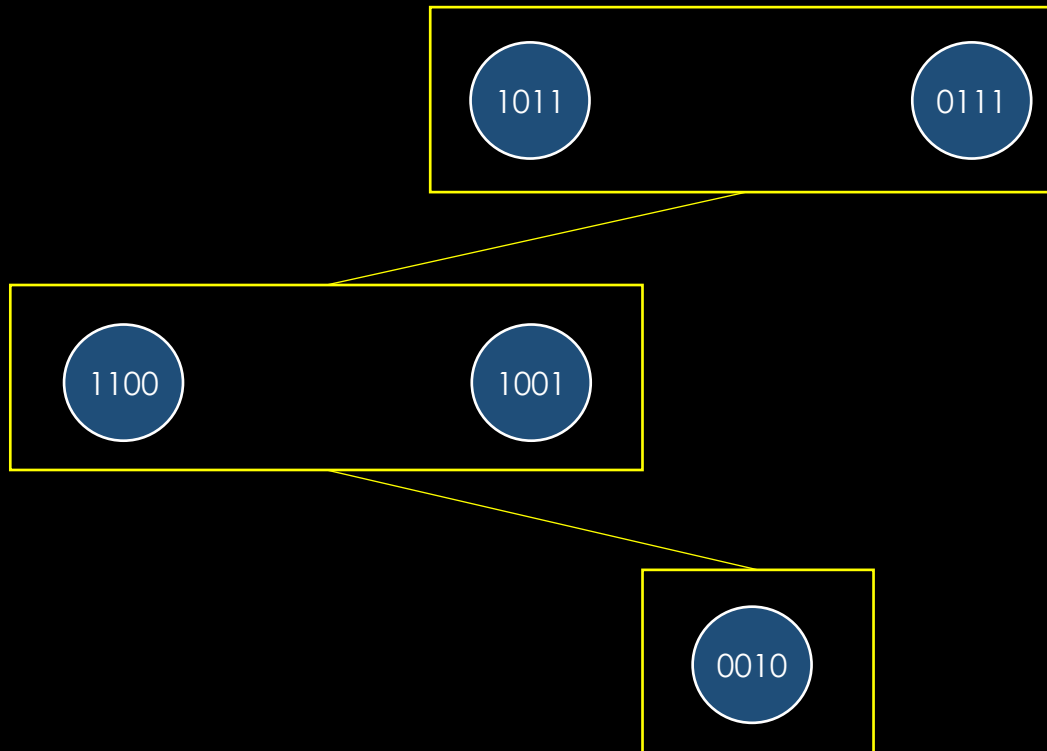
Can we rank them?

Sanitation
Water
Shelter
Health



Collapsing complexity (e.g. by counting 1s)?

Sanitation
Water
Shelter
Health



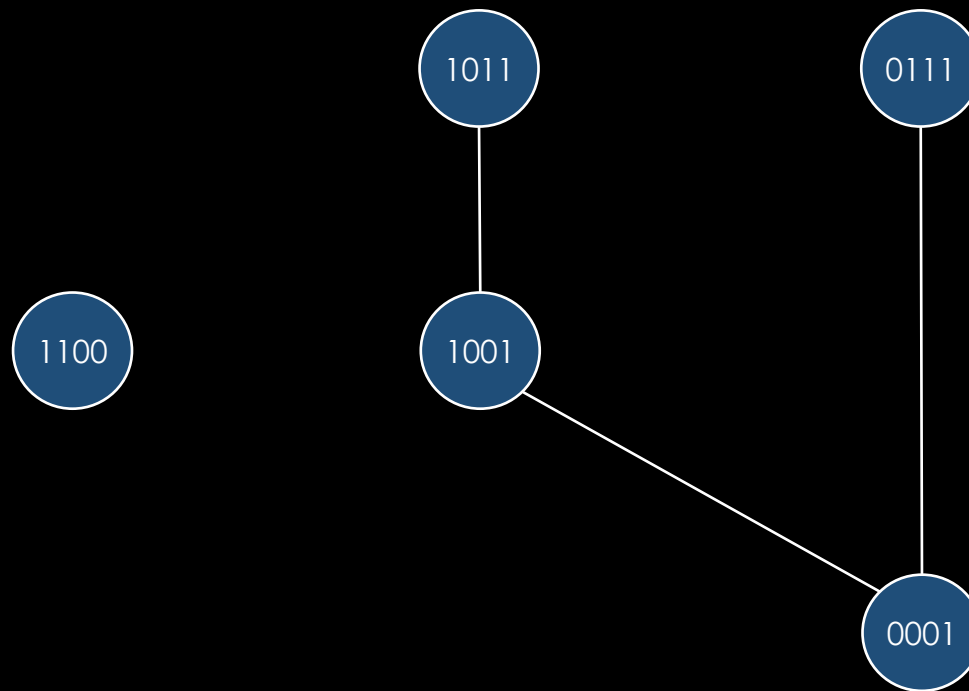
Preserving complexity and the nature of data

DATA: **ORDINAL**

COMPLEXITY: **INCOMPARABILITIES**

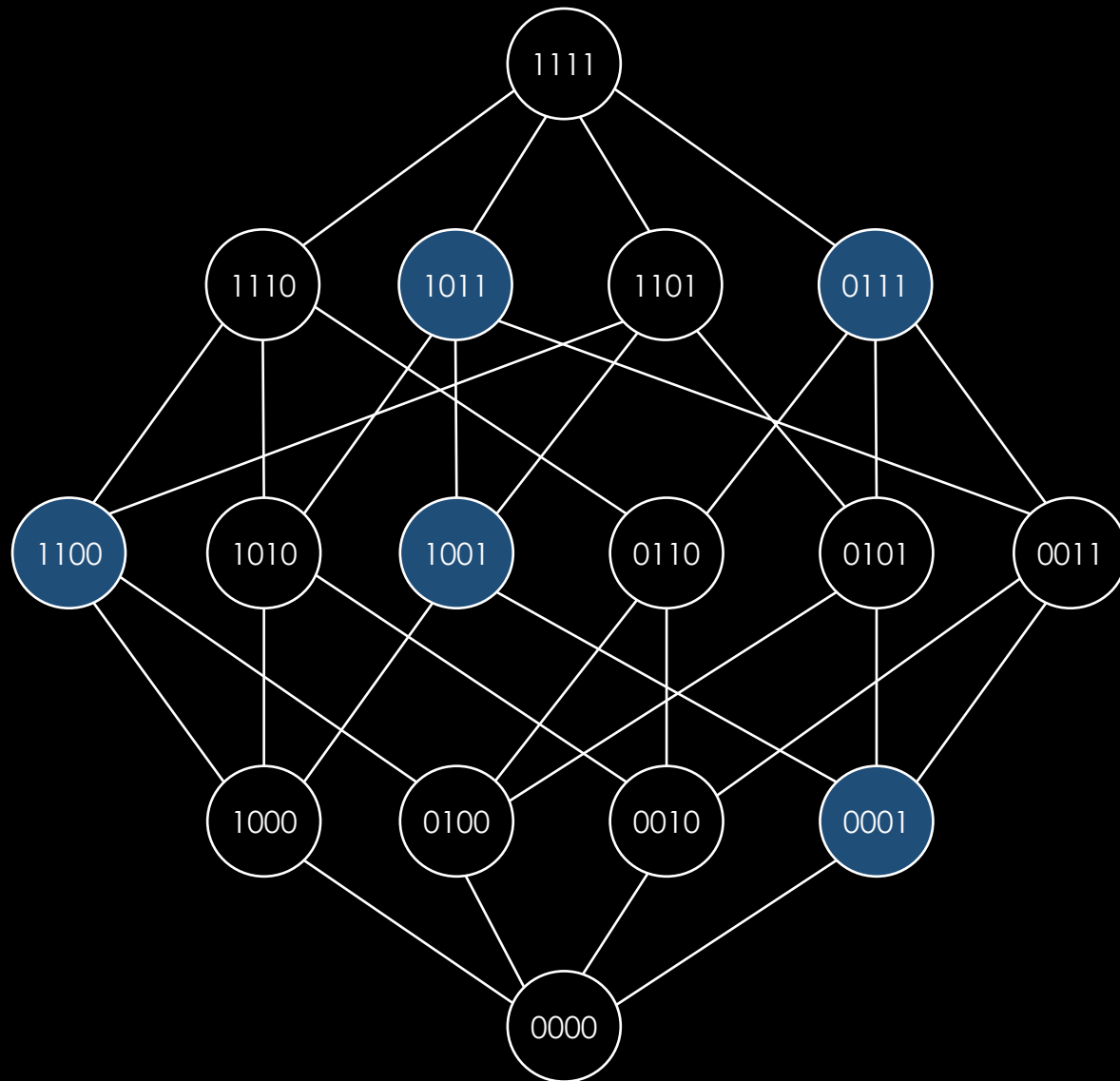
The data structure: partially ordered set

Sanitation
Water
Shelter
Health



The «space» of wellbeing

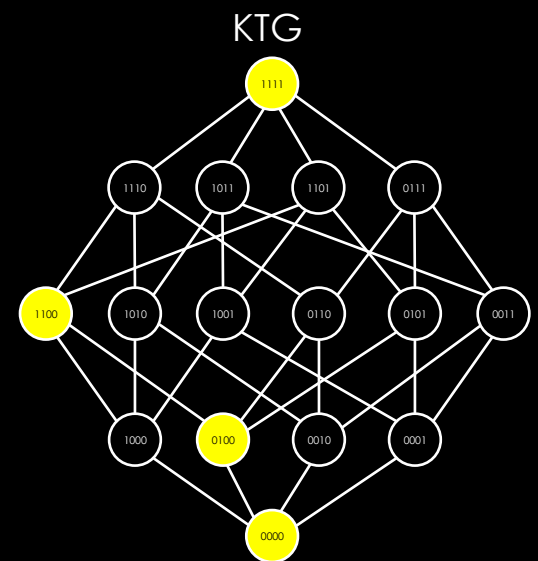
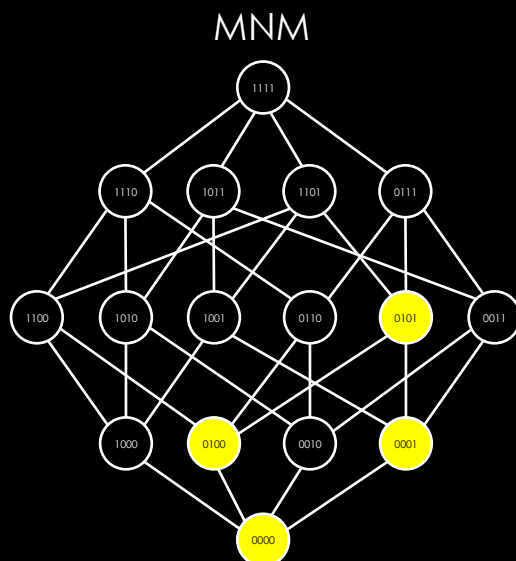
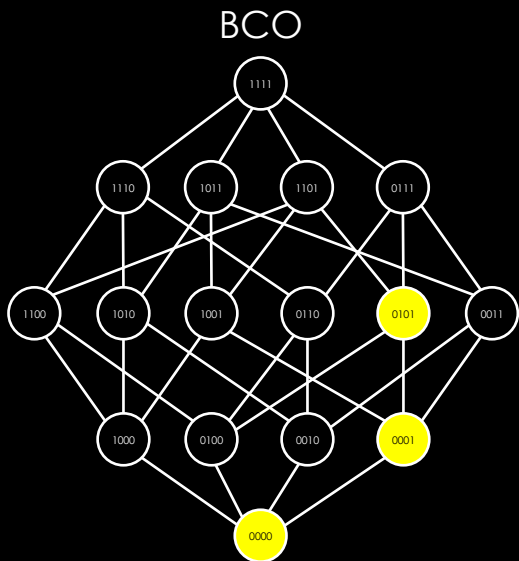
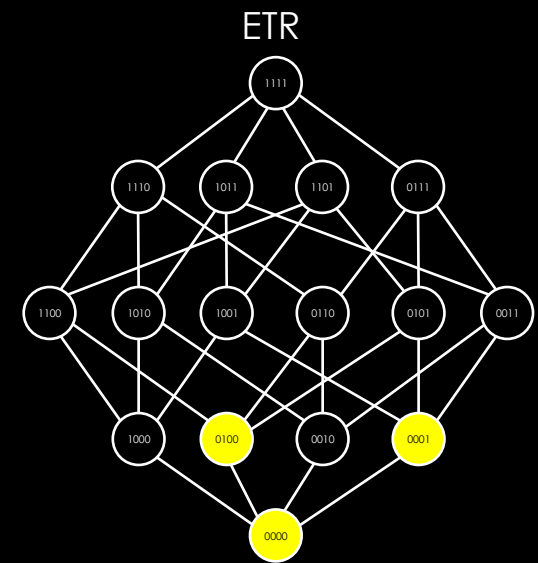
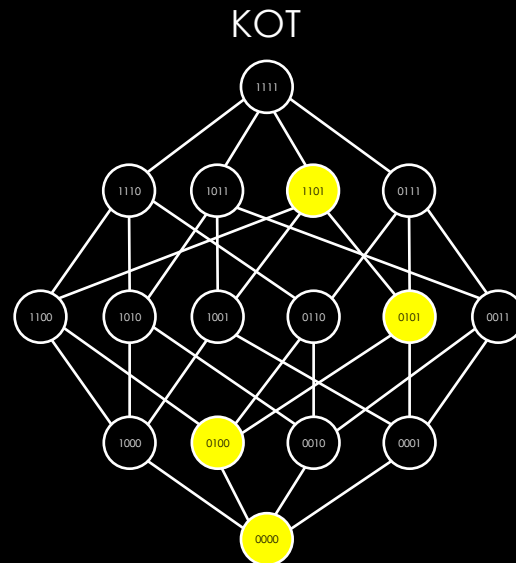
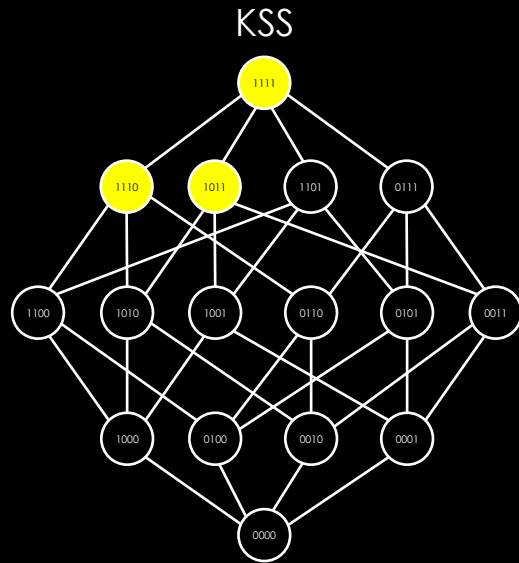
Sanitation
Water
Shelter
Health



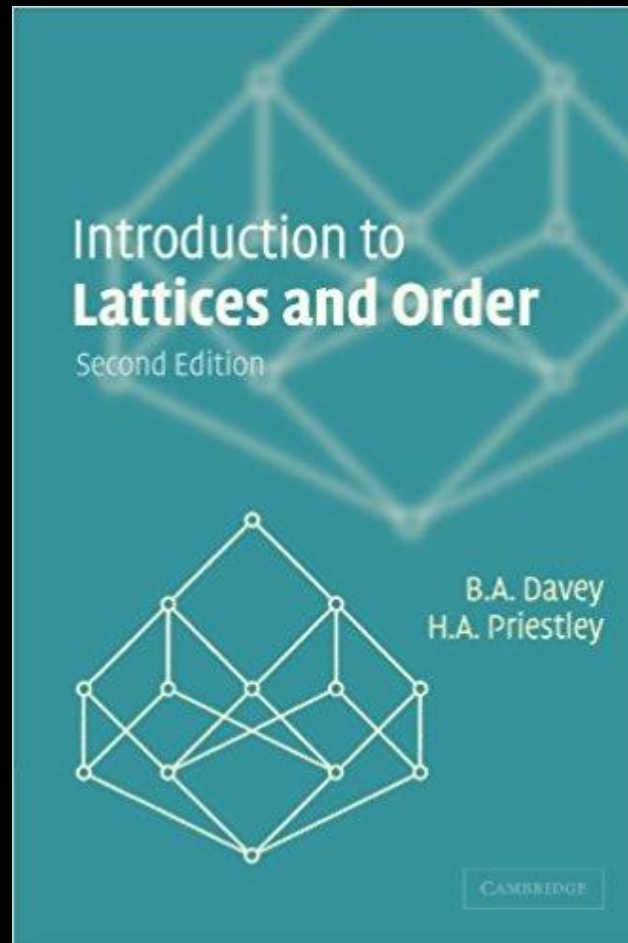
Comparing frequency distributions: regions of Congo

PRF	DRC	KSS	BCO	BDD	ETR	ORT	NKV	MNM	SKV	KTG	KOT	KOC
1111	0.060	0.307	0.059	0.000	0.000	0.010	0.010	0.000	0.070	0.130	0.030	0.000
0111	0.010	0.020	0.020	0.000	0.010	0.000	0.020	0.020	0.030	0.000	0.000	0.010
1011	0.020	0.119	0.079	0.000	0.000	0.000	0.010	0.000	0.010	0.020	0.010	0.010
0011	0.010	0.030	0.020	0.020	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.010
1101	0.040	0.089	0.020	0.000	0.000	0.030	0.010	0.020	0.040	0.080	0.120	0.000
0101	0.060	0.010	0.099	0.020	0.052	0.040	0.040	0.173	0.050	0.070	0.110	0.110
1001	0.030	0.059	0.020	0.010	0.000	0.010	0.101	0.031	0.030	0.060	0.030	0.060
0001	0.150	0.030	0.238	0.404	0.165	0.080	0.071	0.112	0.100	0.060	0.070	0.220
1110	0.030	0.119	0.020	0.000	0.010	0.010	0.030	0.000	0.040	0.070	0.050	0.000
0110	0.010	0.020	0.030	0.000	0.000	0.000	0.020	0.010	0.030	0.000	0.040	0.010
1010	0.020	0.059	0.010	0.000	0.000	0.020	0.030	0.010	0.050	0.020	0.020	0.000
0010	0.030	0.020	0.010	0.111	0.010	0.010	0.010	0.010	0.020	0.010	0.010	0.020
1100	0.040	0.040	0.000	0.000	0.010	0.040	0.010	0.031	0.040	0.100	0.080	0.000
0100	0.140	0.010	0.059	0.071	0.134	0.250	0.081	0.357	0.160	0.100	0.240	0.160
1000	0.040	0.040	0.010	0.020	0.010	0.030	0.152	0.031	0.110	0.060	0.020	0.030
0000	0.310	0.030	0.307	0.343	0.588	0.470	0.404	0.194	0.220	0.220	0.170	0.360

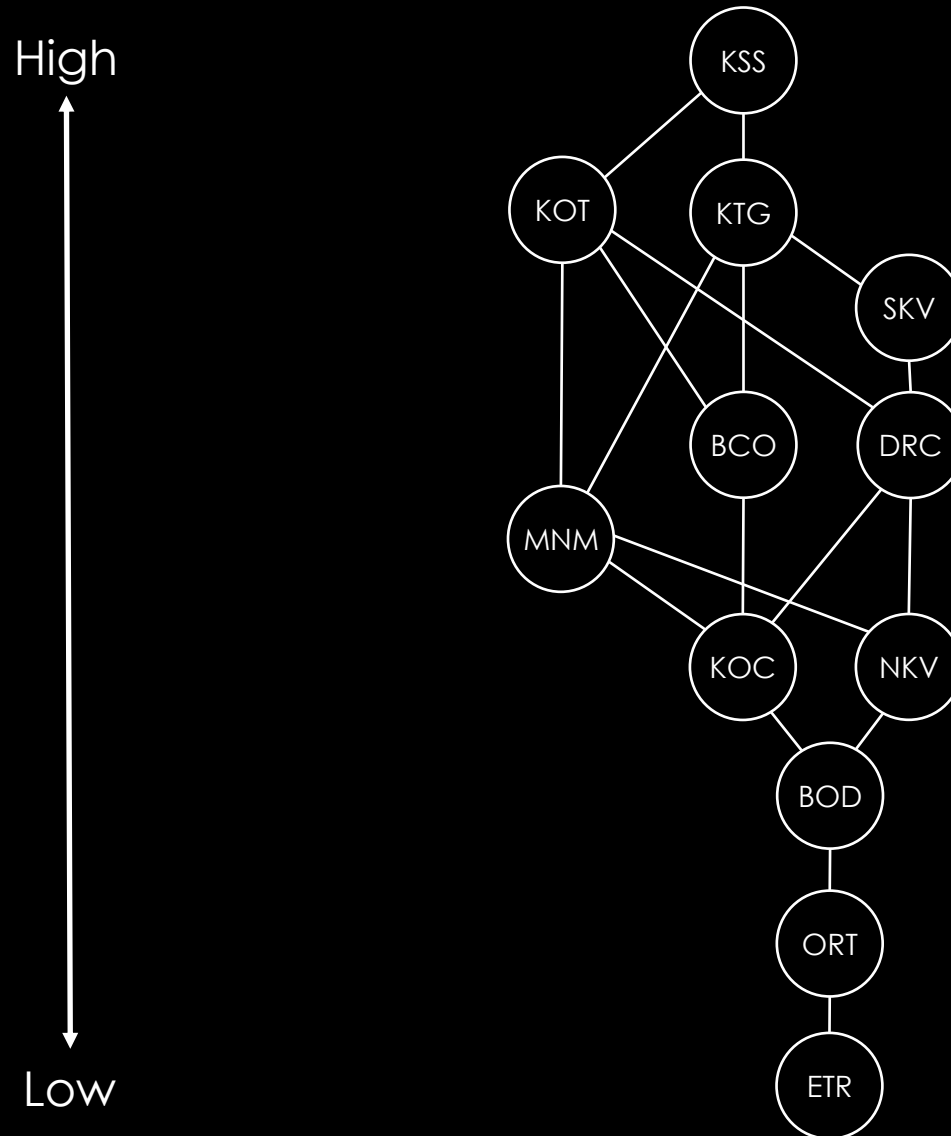
The «space» of wellbeing: different distributions



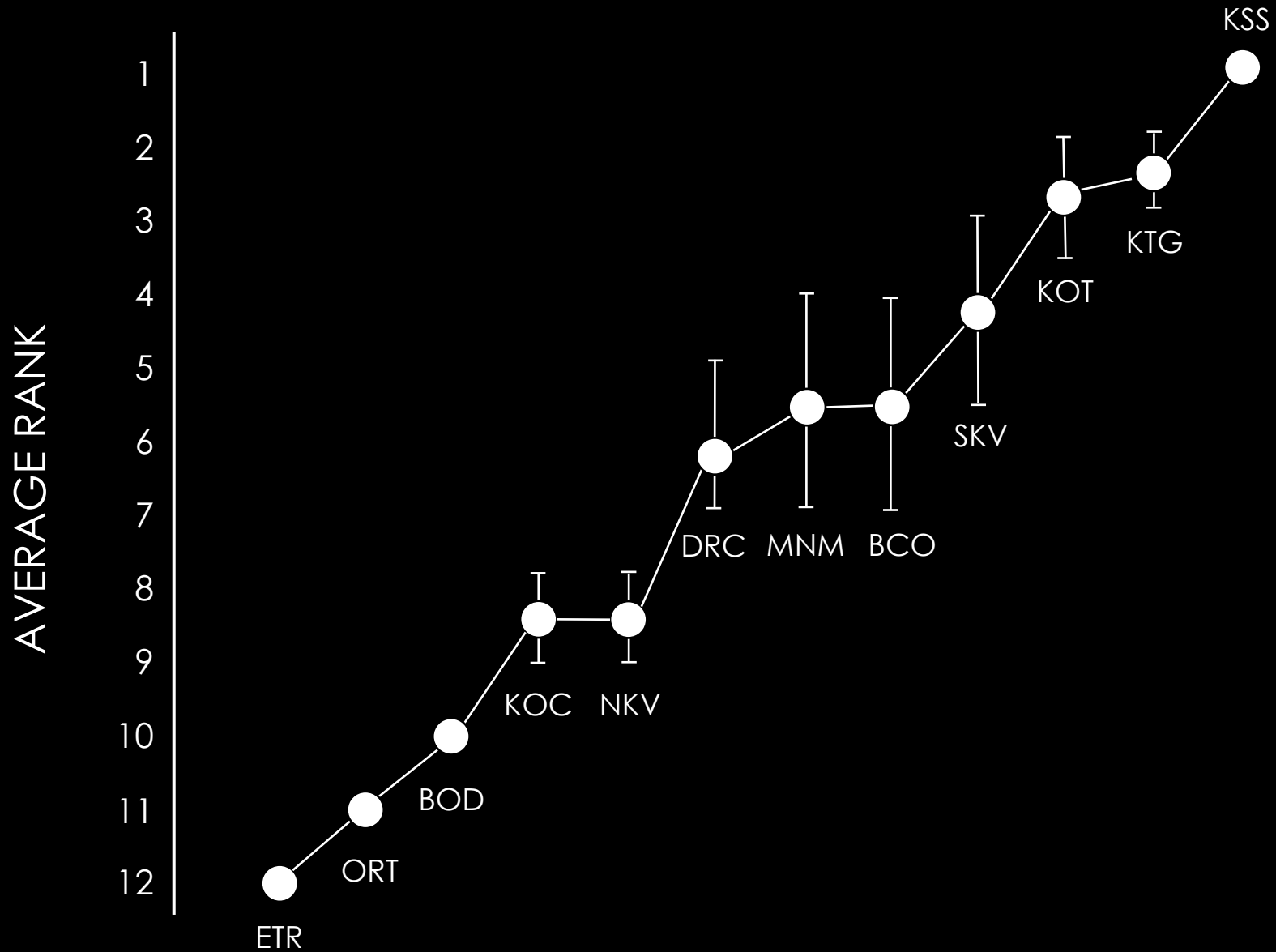
Can we fill the gap, by developing proper procedures to synthesize multi-dimensional systems of ordinal indicators?



Child wellbeing in Congo: regions partial order



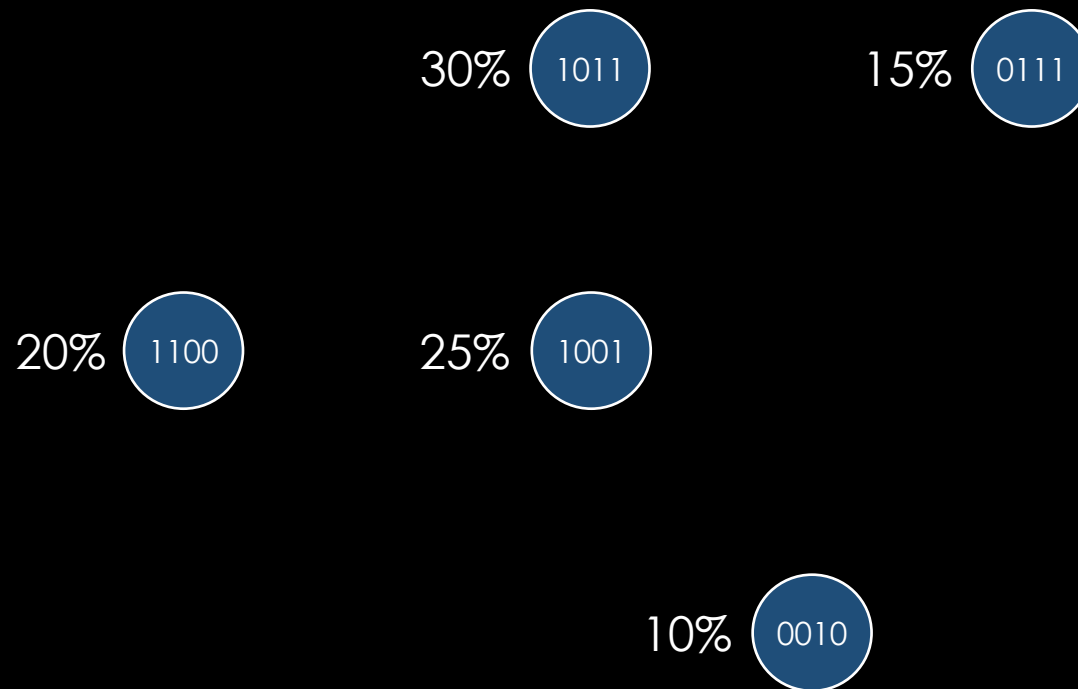
Child wellbeing in Congo: final ranking



Other questions...

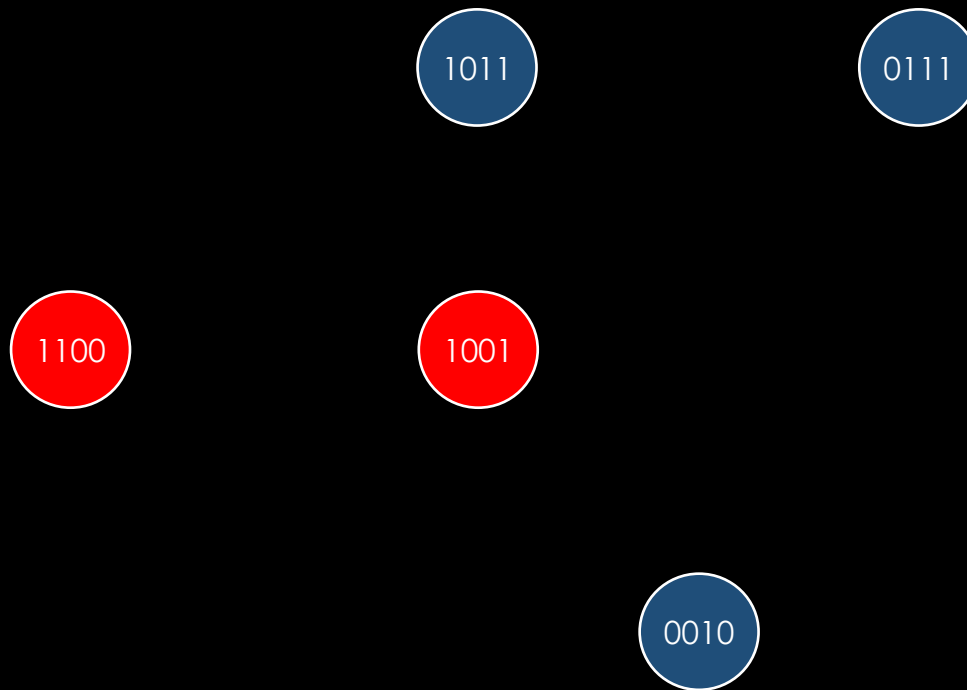
How much unequal is children wellbeing?

Sanitation
Water
Shelter
Health



What is the wellbeing degree?

Sanitation
Water
Shelter
Health




Soc Indic Res

<https://doi.org/10.1007/s11205-018-1856-9>



CrossMark

Multidimensional Analysis of Deprivation and Fragility Patterns of Migrants in Lombardy, Using Partially Ordered Sets and Self-Organizing Maps

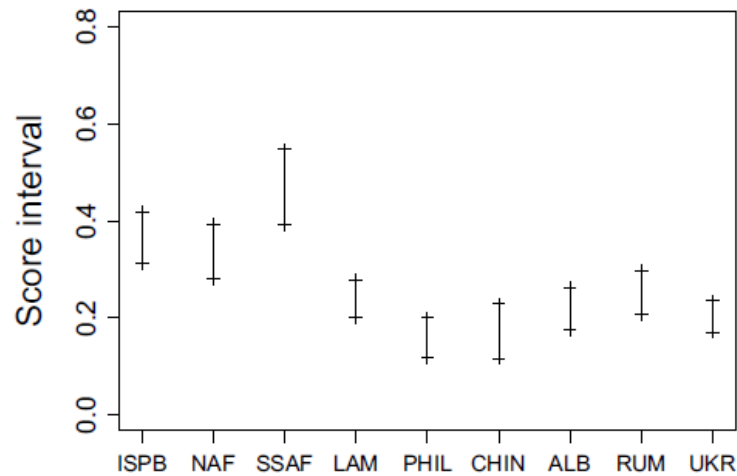
Alberto Arcagni¹  · Elisa Barbiano di Belgiojoso¹ · Marco Fattore¹ · Stefania M. L. Rimoldi¹

Accepted: 5 February 2018

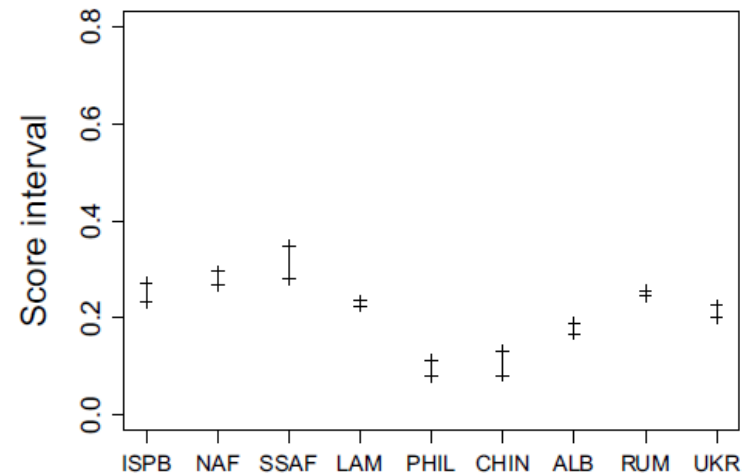
© Springer Science+Business Media B.V., part of Springer Nature 2018

Immigrants deprivation

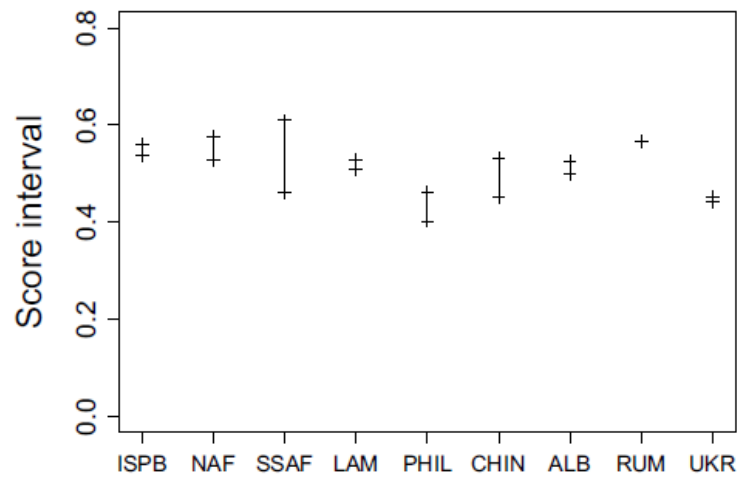
Mean poverty identification score



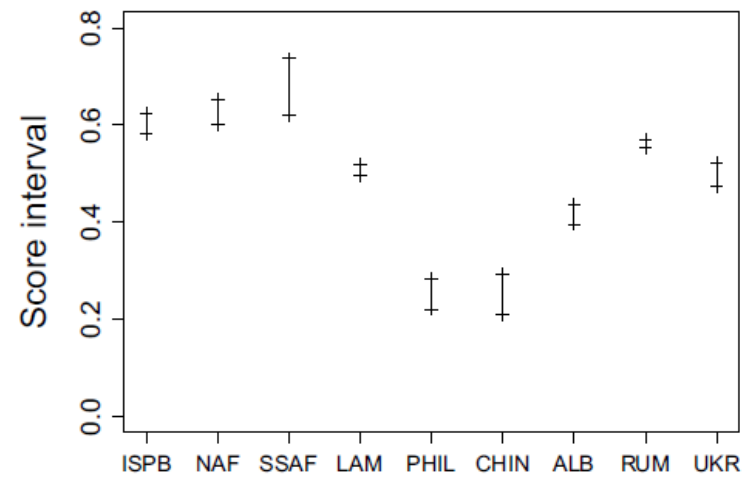
Mean fragility identification score

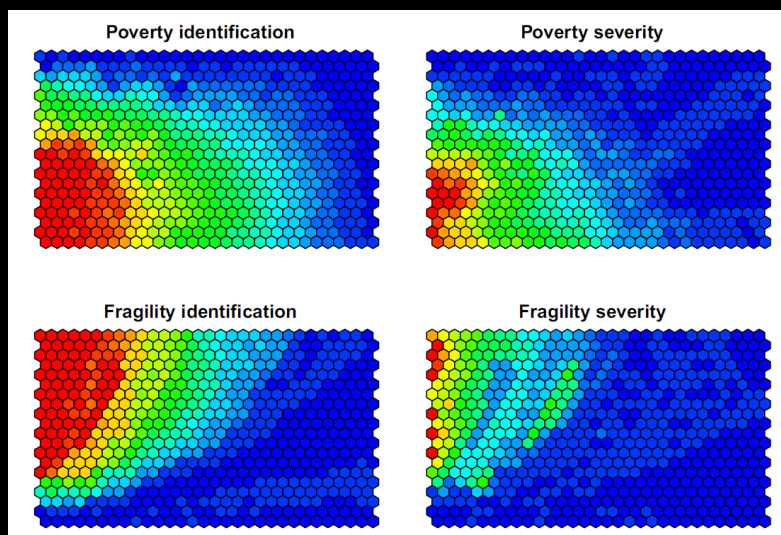


Poverty polarization

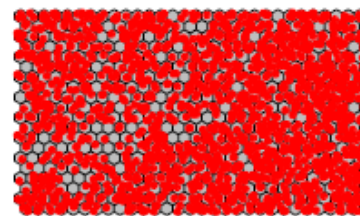


Fragility polarization

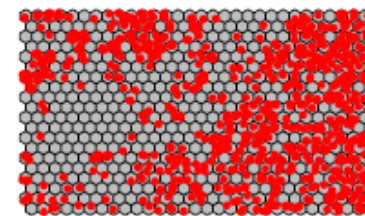




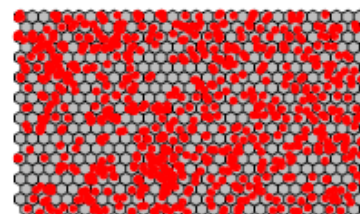
Immigrants distribution on the map



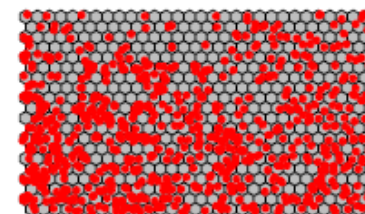
Latin America



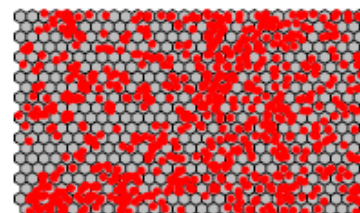
North Africa



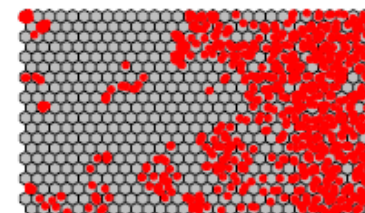
Sub Saharian Africa



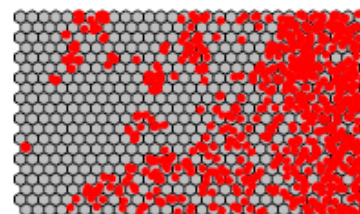
IPSB



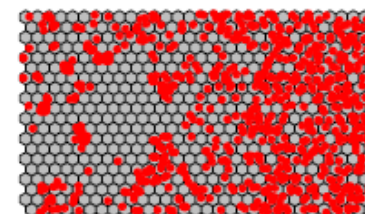
Philippines



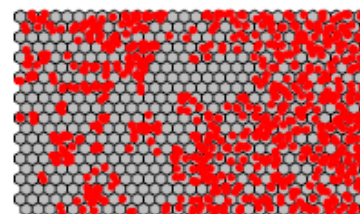
China



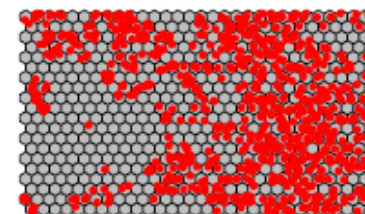
Albania



Romania



Ukraine and Moldova



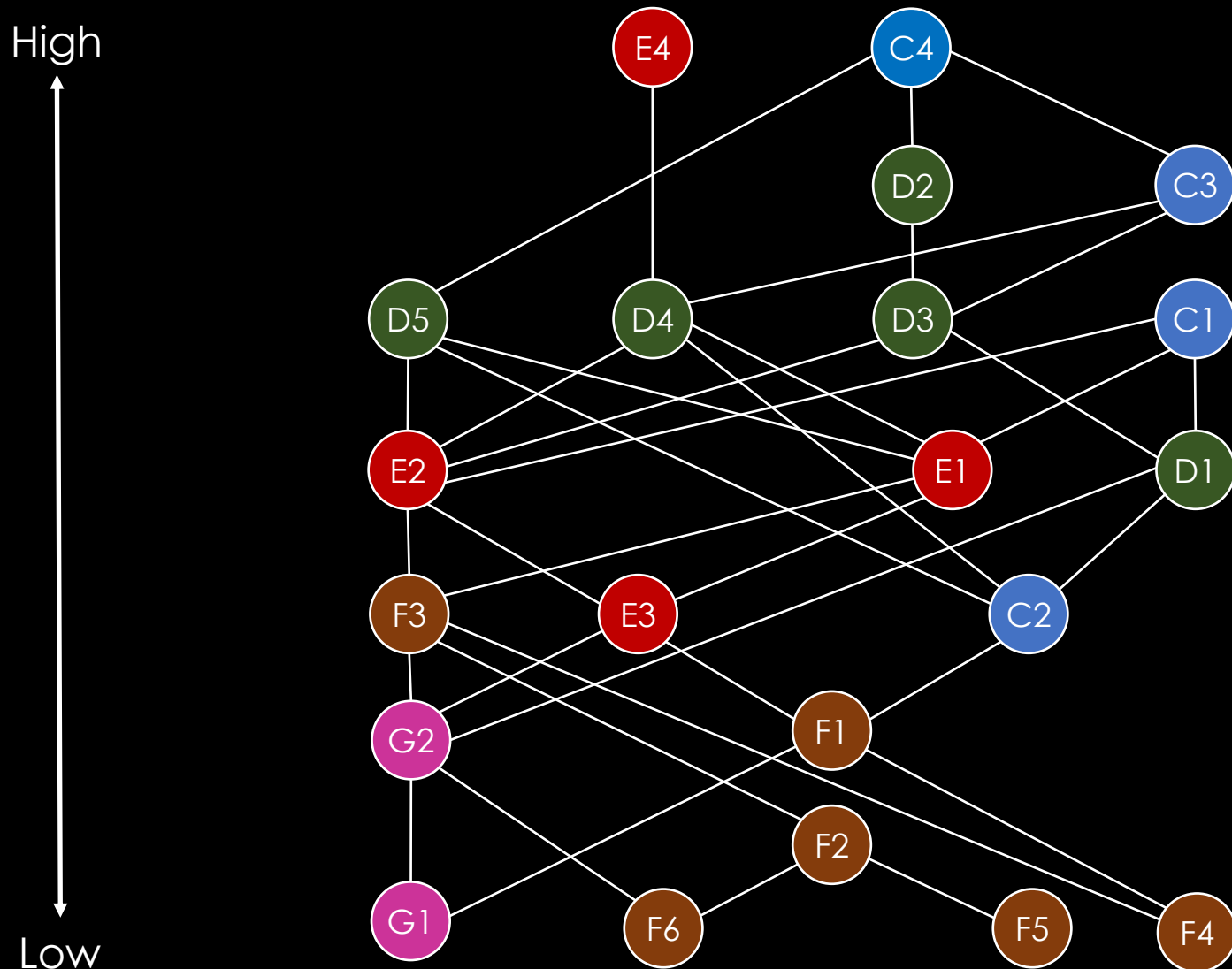
Non-aggregative Assessment of Subjective Well-Being

Marco Fattore, Filomena Maggino, and Alberto Arcagni

Subjective well-being

Regions	ID	Head count ratio			Gap		
		Total	Males	Females	Total	Males	Females
Piemonte - Valle d'Aosta	1	0.16	0.16	0.17	0.13	0.12	0.13
Lombardia	2	0.14	0.13	0.16	0.11	0.11	0.12
Trentino-Alto Adige	3	0.08	0.07	0.08	0.08	0.08	0.08
Veneto	4	0.15	0.14	0.16	0.12	0.11	0.13
Friuli Venezia Giulia	5	0.14	0.13	0.15	0.12	0.10	0.13
Liguria	6	0.14	0.12	0.15	0.11	0.10	0.12
Emilia Romagna	7	0.14	0.13	0.16	0.11	0.11	0.12
Toscana	8	0.15	0.14	0.16	0.12	0.11	0.13
Umbria	9	0.20	0.17	0.23	0.15	0.13	0.17
Marche	10	0.17	0.15	0.18	0.12	0.11	0.13
Lazio	11	0.19	0.17	0.22	0.13	0.12	0.15
Abruzzo	12	0.20	0.17	0.22	0.13	0.12	0.15
Molise	13	0.18	0.18	0.18	0.12	0.12	0.12
Campania	14	0.24	0.22	0.26	0.15	0.14	0.16
Puglia	15	0.24	0.21	0.26	0.16	0.14	0.17
Basilicata	16	0.20	0.16	0.23	0.13	0.11	0.14
Calabria	17	0.22	0.19	0.25	0.15	0.13	0.16
Sicilia	18	0.23	0.21	0.24	0.15	0.15	0.16
Sardegna	19	0.22	0.20	0.24	0.15	0.13	0.17
Italy		0.18	0.16	0.20	0.13	0.12	0.14

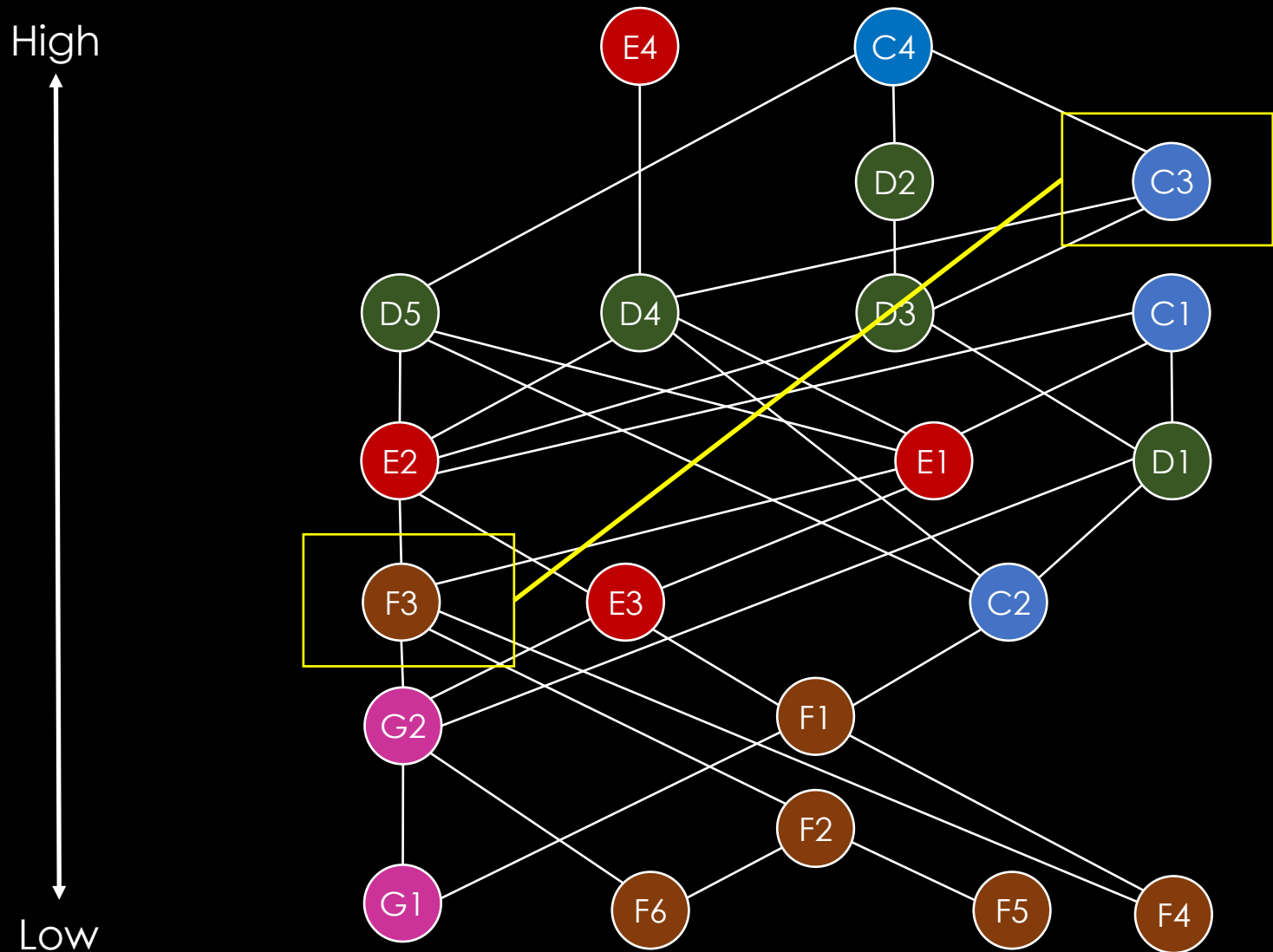
Italian competitiveness by regions (basic, efficiency, innovation)



Italian competitiveness by regions (basic, efficiency, innovation)

Code	Official ranking	Posetic ranking
ITC4	1	1
ITD5	2	6
ITE4	3	2
ITD2	4	4
ITC3	5	3
ITC1	6	5
ITD4	7	7
ITD3	8	8
ITE1	9	10
ITE2	10	11
ITD1	11	9
ITE3	12	13
ITC2	13	12
ITF1	14	15
ITF2	15	17
ITF3	16	14
ITG2	17	16
ITF5	18	21
ITF4	19	18
ITF6	20	20
ITG1	21	19

Which is the gap?)



Partially ordered data and concepts are spreading

Well-being and multidimensional deprivation measurement

Comparison of fiscal systems and policies

Ranking of regions by pollution

Refugee's relocation in UE

Extensions of utility theory

Ranking of regions by multi-dimensional deprivation distributions

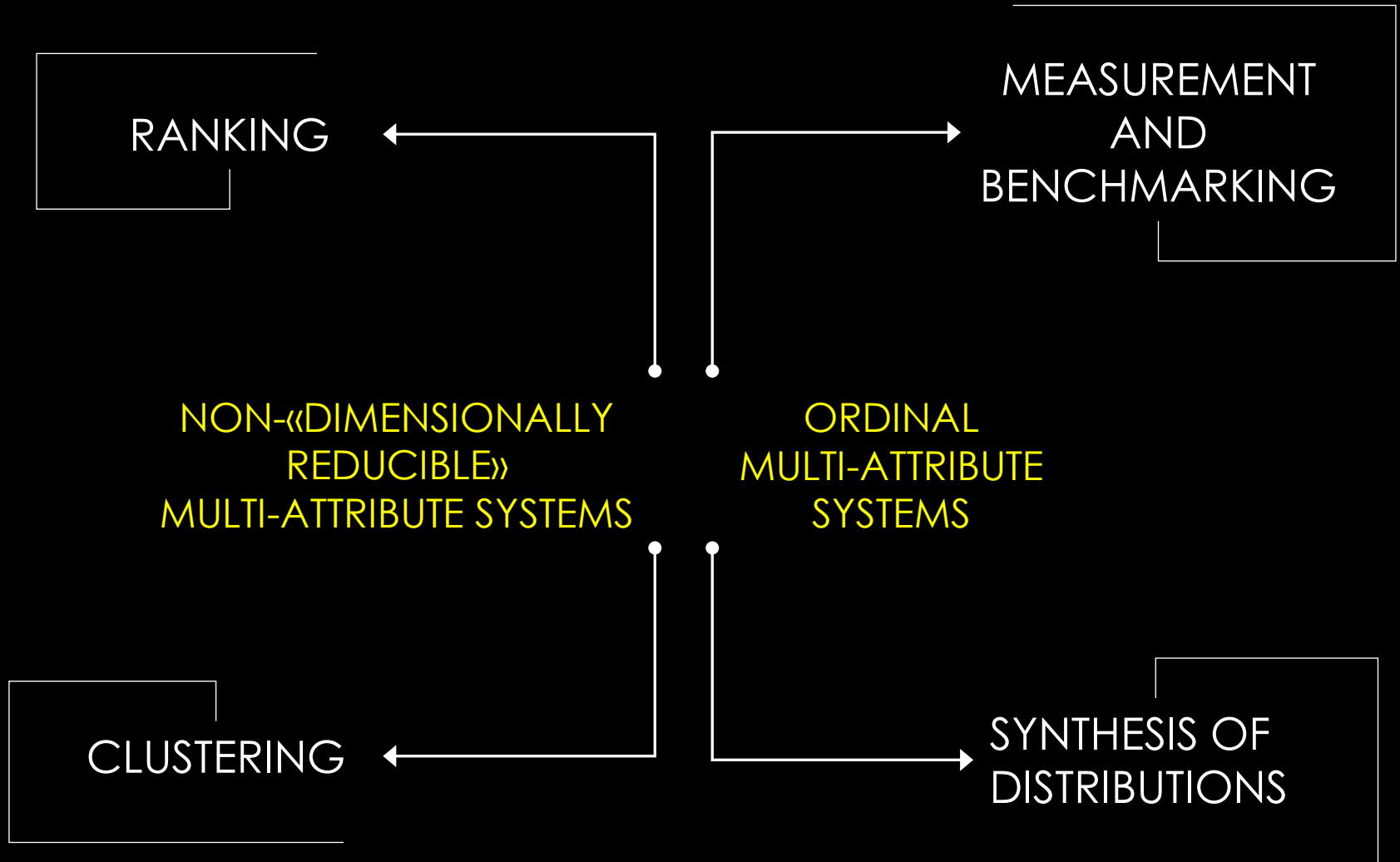
Sustainability

Environmental sciences and ecology

and many others...

General discussion

Purposes and goals



The issue for indicator construction

THE UNDERLYING TRAIT IS
INHERENTLY COMPLEX

DATA CANNOT BE TREATED
BY STANDARD MATH TOOLS

NON-«DIMENSIONALLY
REDUCIBLE»
MULTI-INDICATOR SYSTEMS

ORDINAL
MULTI-INDICATOR
SYSTEMS

DISTANCES, ANGLES,
DIFFERENCES, RATIOS,
SUMS, AVERAGES...

COMPARABILITY
INCOMPARABILITY
< = >

NO

YES

Topic	Example	SW implementation
Ranking	Competitiveness	Implemented
Benchmarking	Multi-dim. poverty	Implemented
Inequality	Well-being	Implemented (basic form)
Distribution comparisons	Well-being	Implemented

What are we about to develop?

Topic	What
-------	------

Ranking	Extension to hierarchical MISes
---------	---------------------------------

Inequality	Decomposition
------------	---------------

Cluster	To be developed from scratch
---------	------------------------------

SW implementation	To be improved
-------------------	----------------

Towards the scientific community

