Human Development Research Paper 2010/39

Graphical Statistical Methods for the Representation of the Human Development Index and its Components

César A. Hidalgo

United Nations Development Programme
Human Development Reports
Research Paper

September 2010

Human Development
Research Paper
2010/39

Graphical Statistical Methods for the Representation of the Human Development Index and its Components

César A. Hidalgo

United Nations Development Programme
Human Development Reports
Research Paper 2010/39
September 2010

Graphical Statistical Methods for the Representation of the Human Development Index and its Components

César A. Hidalgo

César A. Hidalgo is Assistant Professor at the Massachusetts Institute of Technology Media Laboratory and head of the Macro Connections Research Group. He is also Faculty Associate at Harvard's Centre for International Development. E-mail: hidalgo@mit.edu.

Comments should be addressed by email to the author.

Abstract

In this paper we introduce five graphical statistical methods to compare countries level of development relative to other countries and across time. For this, we use seven panels of data on the Human Development Index and its components, containing information on more than 100 countries for more than 35 years. We create visual comparisons of the level of development of countries relative to each other, and across time, through five different visualization techniques: (i) Rankings (ii) Values (iii) Distributions (iv) visual metaphors (*The Development Tree)*, and (v) networks, by introducing the concepts of Partial Ordering Networks (PON) and Development Reference Groups (DRG).

The graphical exploration of both, values and distributions, show a saturation of both the education and life dimensions of the HDI, suggesting a need to extend the definitions of this components to include either more subcomponents, or completely new measures that could help differentiate between countries facing different development challenges. The Development Tree and the Partial Ordering Network, on the other hand, are used to create graphical narratives of countries and regions. The simplicity of the Development Tree makes it an ideal graphical metaphor for branding the HDI in a multilingual setting, whereas Partial Ordering Networks provide a more organic way to group countries according to their levels of development and connect countries to those with similar development challenges. We conclude by arguing that graphical statistical methods could be used to help communicate complex data and concepts through universal cognitive channels that are heretofore underused in the development literature.

Keywords: Human Development Index, Visualization, The Development Tree, Partial Ordering

JEL classification: B4, C82, O11

INTRODUCTION

The Origins of Graphical Statistical Methods (Even charts had to be invented)

Joseph Priestley was an English theologian, clergyman and natural philosopher that lived a prosperous academic life in the 18th century. Among Priestley's most noteworthy contributions was the isolation of oxygen in its gaseous state, a honor that he shared together with Carl Wilhelm Scheele and Antoine Lavoisier. The interests of Priestley, however, extended far beyond chemistry and the Natural Sciences. Indeed, Joseph Priestley was an avid historian and educator who wrote several books on world history, politics and its defining characters, including his "lectures on history and general policy", in which he summarized many of his ideas in forms of government, finance, taxation, and colonies, among other topics. Some of the most outstanding contributions of this polymath, however, came not in the form of writing, but rather as illustrations. In 1765, Joseph Priestley created "A Chart of Biography", a portion of which appears in Figure 1. This chart is believed to be one of the earliest timelines and was created by Priestley to illustrate the contemporaneity of ancient philosophers and statesman. Priestley believed that charts, such as this one could help summarize and clarify important aspects of history making them a useful teaching and analytical resource and devoted an important part of his life to the creation of charts of this sort.

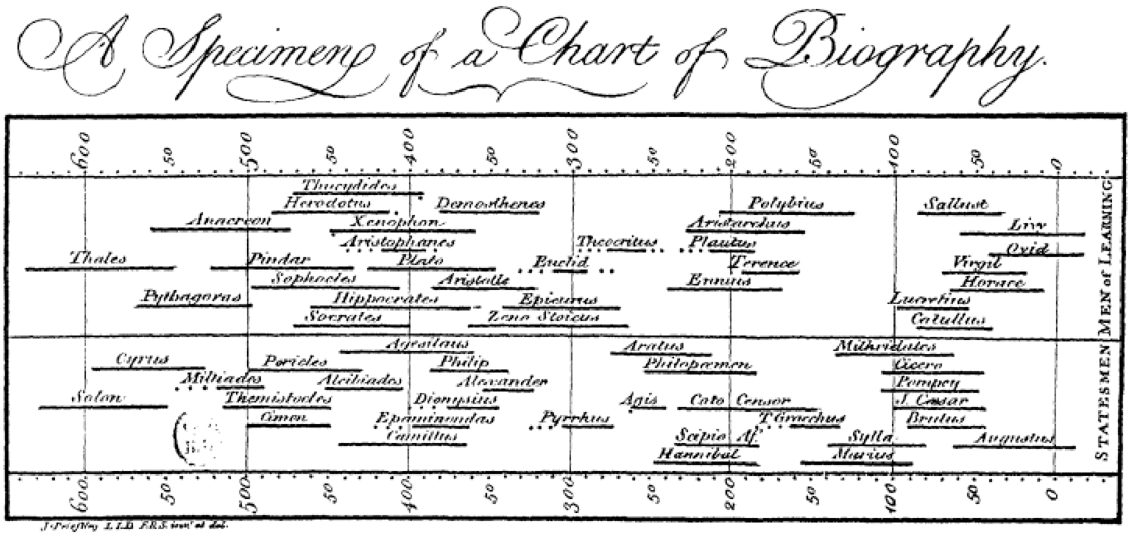

Figure 1. An extract from a chart of biography, by Joseph Priestley (1765). The chart indicates with lines the lives of important statesmen and philosophers that lived between 650 BC and 50AC.

Joseph Priestly used visualizations not only to illustrate the lives of important individuals, but also to summarize the dominance and breadth of civilizations. His 1769 work: "A New Chart of History" (Figure 2), illustrates the temporal and spatial extent of different civilizations and empires. One of the striking features of this visualization is the patterns that emerge from different historical periods. Priestley's Chart of History clearly shows that while some historical periods were characterized by large empires, other periods were marked by a substantial degree of political fragmentation. The chart can also be interpreted as a measure of the spatial and temporal reach of different civilizations, with the area occupied by each one of these cultures as a proxy for their power. The graphical nature of this chart of history makes it easy to make quick comparisons regarding the duration and geographical extent of different civilizations, making it an ideal teaching tool for historians and educators.

A younger, and more statistically oriented contemporary of Joseph Priestley, was William Playfair. Playfair was a Scottish engineer that some considered to be the father of graphical statistical methods, such as the bar chart and the pie chart. Playfair, who once worked as an assistant to his famous fellow Scotsman James Watts, believed that images were able to explain concepts and ideas better than words. This is well exemplified in some of its most well known images. Figure 3 shows two of Playfair's figures, one showing a Pie Chart used to compare the fractions of the Turkish Empire that lay in Asia, Africa and Europe, and the other showing the temporal evolution of England's trade balance with Norway and Denmark, which is represented by the shaded area between the curves. While both of these figures seem extremely simple from a modern perspective, they were ground breaking for their time and helped cement the path to a more widespread use of graphical statistical methods in the following centuries.

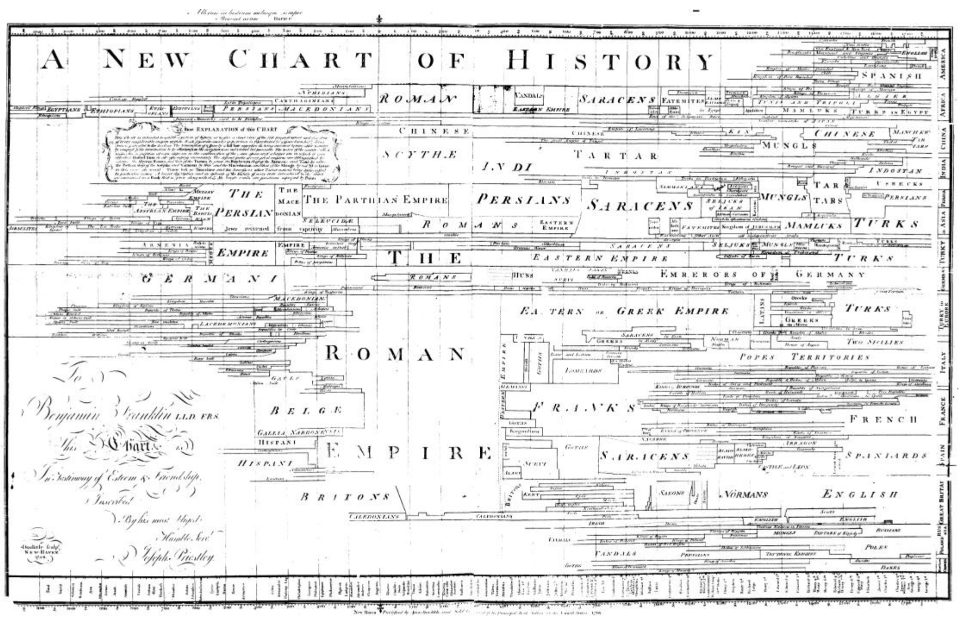

Figure 2. A new chart of History by Joseph Priestley (1769). The chart illustrates, and compare, the duration and geographical extent of mayor historical civilizations.

The 19th century saw a great increase in the utility and quality of visualizations used to describe historical events and socio demographic patterns. Among the most salient of these are the works created by Charles Joseph Minard, Florence Nightingale and John Snow. Minard was a French mathematician and engineer who created several visualizations illustrating historical war campaigns and macroeconomic patterns. The most famous of Minard's work is probably his "*Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813*" a flow map published in 1869 describing Napoleon's Russian Campaign (Figure 4) (Tufte 2006).

Other extremely noteworthy contributions by Minard are related to economic geography and demographics. Here we show two examples of Minard's work in which he uses data on emigration for the year 1858 to show the flow of emigrants from all major continents (Figure 5). Minard also created several visualizations illustrating economic geography patterns, such as the capacity of ports, the production of cattle, the transportation of wine, the exports and imports of cotton from Britain and the export of Britain's coal, depicted in Figure 6 for the year 1864.

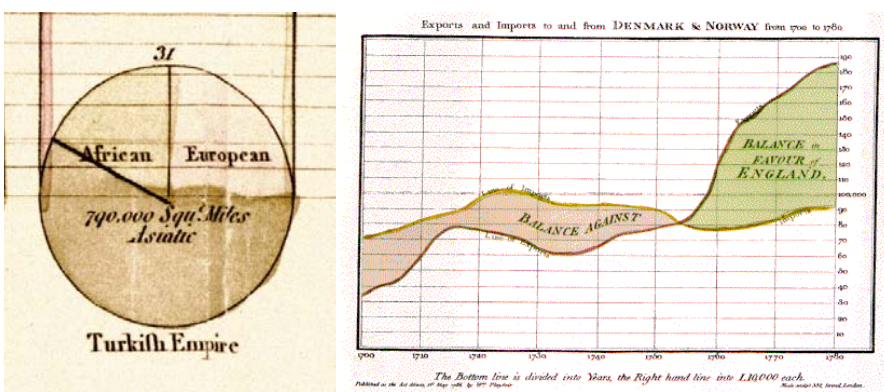

Figure 3 Two examples of William Playfair's work. On the left, one of the earliest Pie Charts illustrating the fractions of the Turkish Empire that lied in Europe, Africa and

Asia in 1789 (created in 1801). And on the right an illustration of the trade balance between England, Denmark and Norway (1786)

While the works of Minard are aesthetically impressive and clearly ahead of his time, the more modest visualization works done by Florence Nightingale and John Snow had the power to change public health for years to come.

Soon after the Crimean war Florence Nightingale was asked by the queen to write a report on the poor sanitary conditions faced by the British Army and to make some recommendations for possible health reforms. After the government did not follow up on her report Nightingale decided to write and submit an annex where she included her now famous Nightingale's rose (Figure 7). Nightingale's Rose is a visualization that the famous nurse used to show that most individuals that died during the Crimean war did so from preventable, rather than catastrophic causes (preventable causes are represented by the light blue area of Figure 7). It is believed that Nightingale thought that the queen will probably not read her report but that her attention could be more easily caught by a picture, rather than a paragraph. Some historian believe that Nightingale's rose ultimately help persuade the government to institute some of the reforms she suggested (Brasseur 2005).

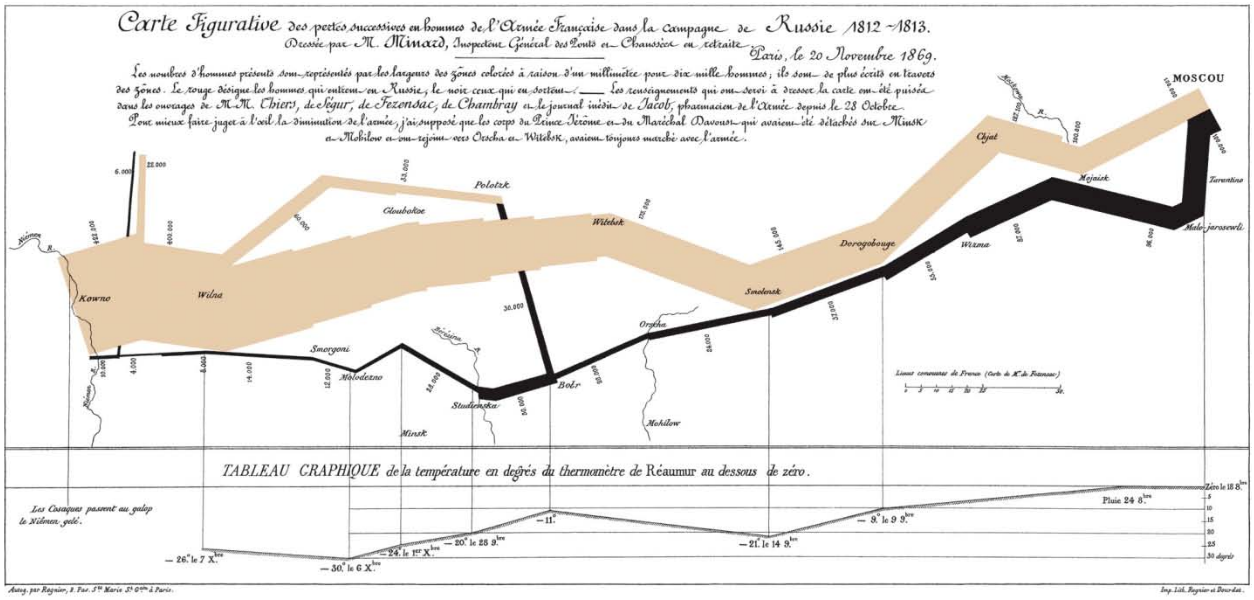

Figure 4 "*Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813*" a flow map published in 1869 describing Napoleon's Russian Campaign by Charles Joseph Minard.

Another 19 century visualization that changed the history of public health was John Snow's Broad street map (Hempel 2007). In 1854 a large cholera outbreak struck the Soho neighborhood in London. To help elucidate the origin of the cholera outbreak, John Snow geo-located the number of cholera infected individuals in a city map, using small horizontal lines to indicate the number of infected subjects in each residence. This map was used to demonstrate that the Cholera outbreak affecting London's Soho district had a well defined source, which was the water pump on Broad Street. Snow's Broad street map is considered by some epidemiological historians as the essential piece of evidence that help debunk the humors theory of disease and opened the ground for the now prevalent germ based theory of infectious diseases, changing medicine and the world forever after.

The works of Priestley, Playfair, Minard, Nightingale and Snow illustrate how visualizations can be used to efficiently represent complex phenomena and convey powerful ideas. Moreover, they show how visualizations can be used as evidence that can ultimately help demonstrate and communicate important aspects of complex phenomena helping people revise their opinion and promote positive change in society, a fact that is specially salient in the world changing works of Nightingale and Snow.

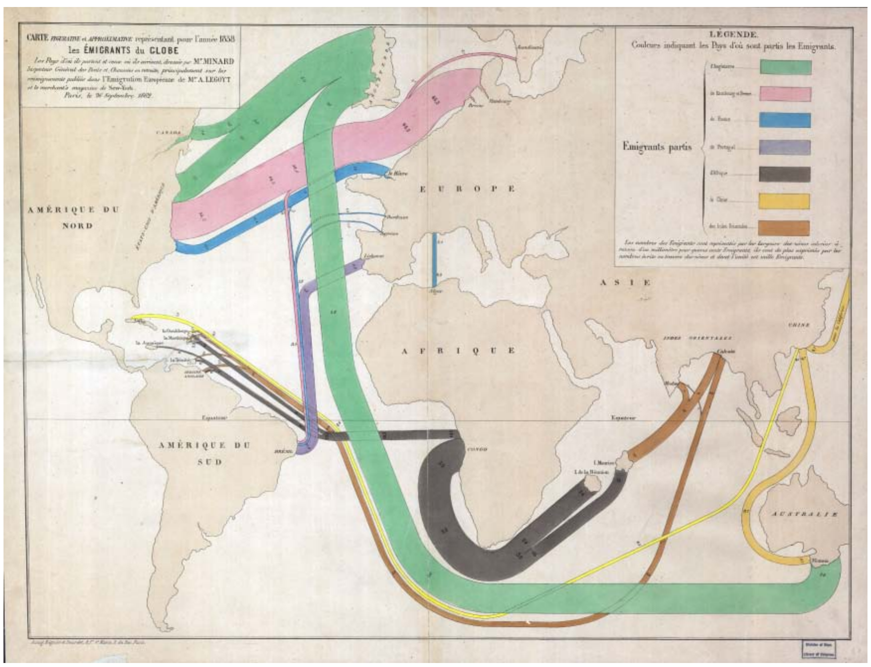

Figure 5 Emigration patterns for the year 1858 by Charles Joseph Minard (1862)

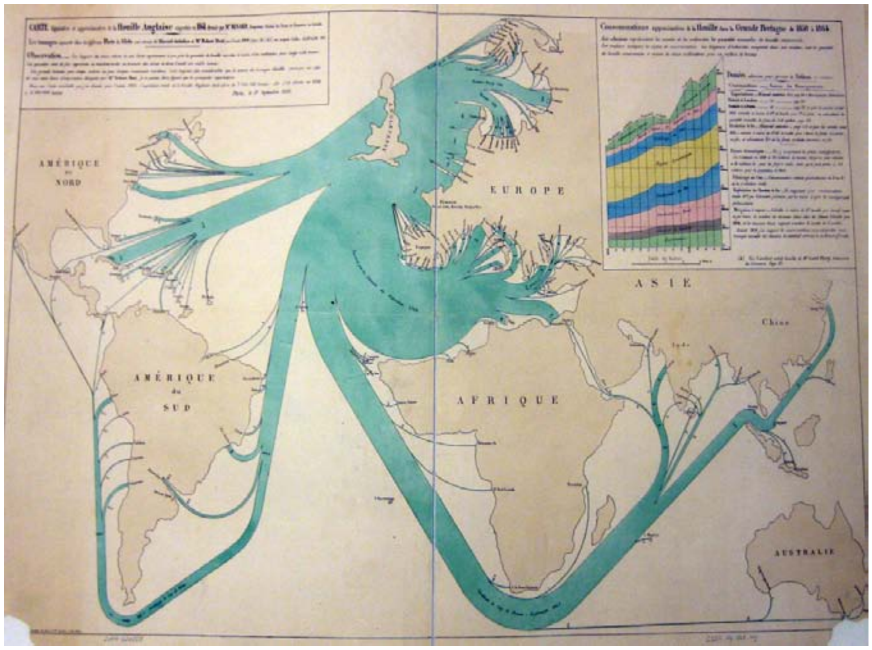

Figure 6 Britain coal exports for the year 1864. Each milimiter of thickness represents 20,000 tons of coal (Minard 1866).

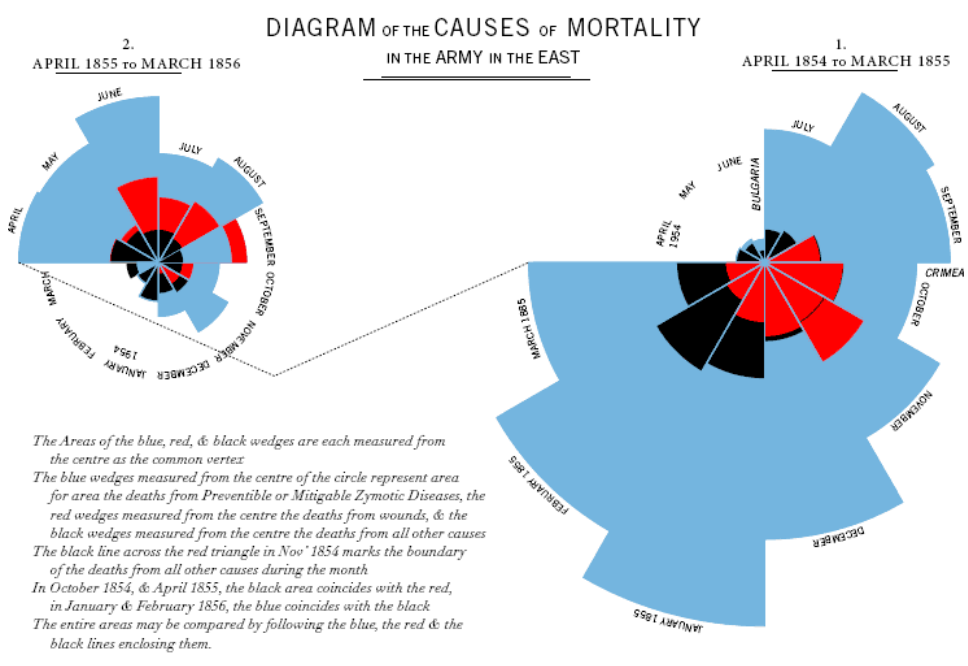

Figure 7 Nightingale's Rose 1858

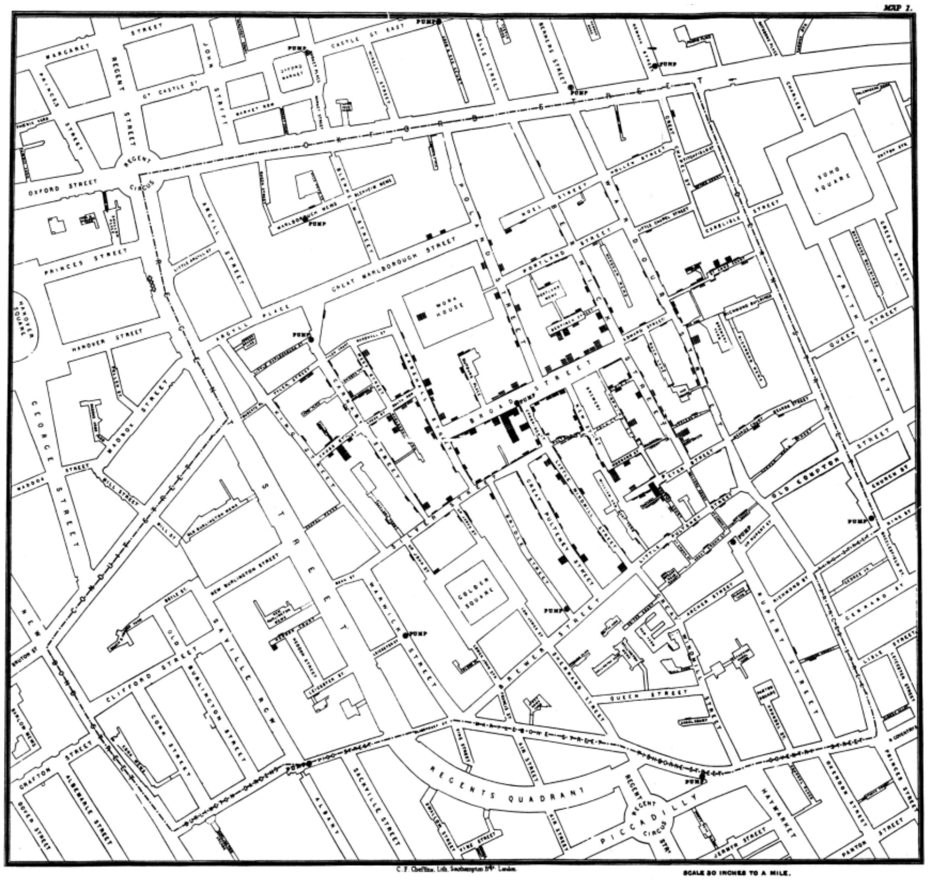

Figure 8 John Snow's Broad Street Map.

MODERN USES OF GRAPHICAL STATISTICAL METHODS

During the last decades improvements in information technology have come together with an increase in the number of graphical statistical methods used by analysts to explore complex social, physical, biological and economic phenomena. Network visualizations are among the most useful of these methods, since they allow analysts to work with data in addition to demographic and geographic layers of information.

A good example of a new graphical statistical method, as well as an analytical concept, that is now used to study the level of industrial development of countries and its evolution is the

Product Space. Figure 9 shows a visualization of the Product Space in which the relationships between more than 700 product categories is illustrated. In this visualization goods are represented by nodes and links connect goods that tend to be exported by the same countries. Here, the co-exporting of goods by countries is taken as a proxy for the likelihood that goods require a similar set of non-tradable inputs or capabilities (Hidalgo et al. 2007). Hence, The Product Space represents a "map" of industrial development that can be used to see the position in the industrial space of a country together with its set of possible industrial options.

Figure 10 uses black squares to visualize the evolution of the export baskets of Malaysia, China and Mexico, illustrating that countries tend to move towards products that are close by to the ones they already produce in the Product Space. The visualization shows that that industrial development is a process constrained by the structure of this network, a fact that is relatively easy to understand graphically, but is complex to illustrate otherwise.

While Figure 9 and Figure 10 show instances of the Product Space using that in which products are disaggregated into 775 categories, Figure 11 shows a more disaggregate visualization of the Product Space containing more than 5,000 different categories. If we approximate the information in Figure 11 as the total number of nodes, links and categories illustrated there, we would conclude that Figure 11 is showing us a representation composed of more than 30,000 values, an amount of data which is equivalent to more than 34 HDR annex pages[1]. Yet, such an interpretation of the volume of information represented in Figure 11 would not be considering the patterns described by these large set of data, an aspect of the pattern of

[1] Considering the annex of the 2009 edition of the Human Development Report.

relatedness between products that is clearly visible from Figure 11 but that could not be extracted from numerical tables.

In the next section we explore different ways to represent time series data on the Human Development Index and its components. The goal of these representations is to show how graphical methods can be used to (i) compare changes in the level of development experienced by countries (ii) make it easier to understand how these changes are tied to each one of the components of the Human Development Index (iii) understand the evolution of the distribution of countries according to HDI and its components and (iv) teach and create awareness about human development by using iconographic representations that can be used to graphically narrate the story of countries and regions.

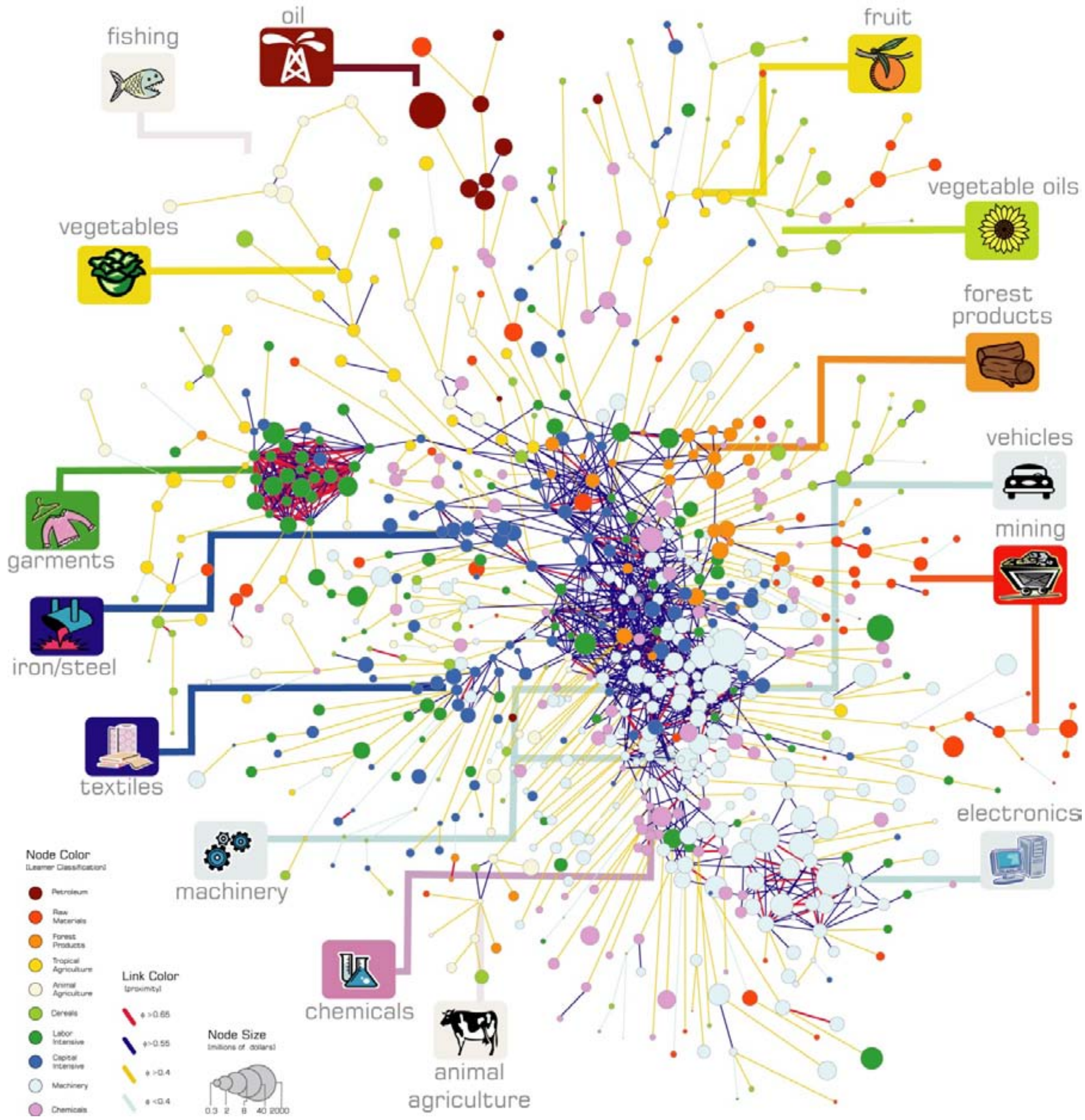

Figure 9 The Product Space for the 1998-2000 period.(Hidalgo et al 2007)

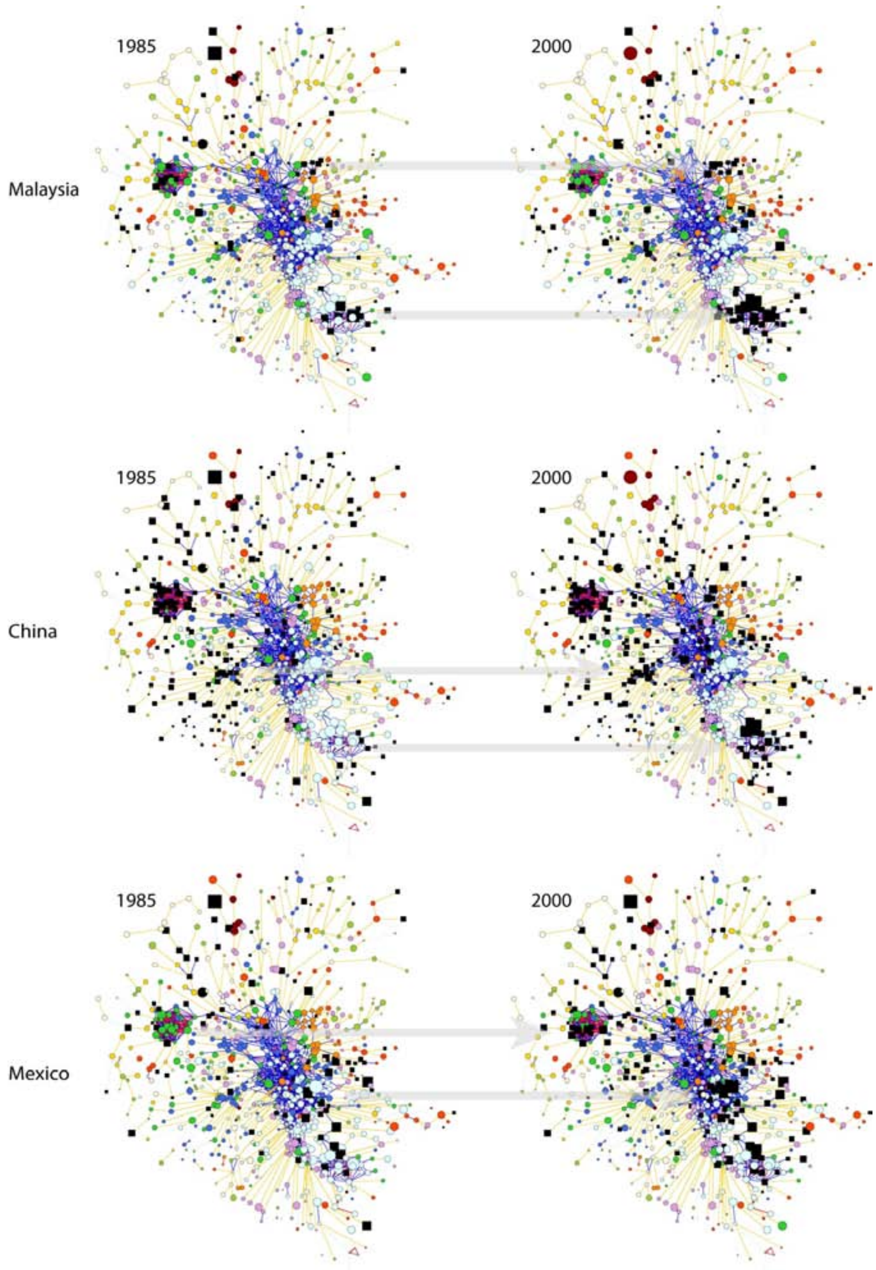

Figure 10 Changes in patterns of comparative advantage for Malaysia, China and Mexico between 1985 and 2000. Products for which these countries have revealed comparative advantage (meaning they are being exported in sufficiently large amounts) are indicated in black squares. (Hidalgo et al 2007)

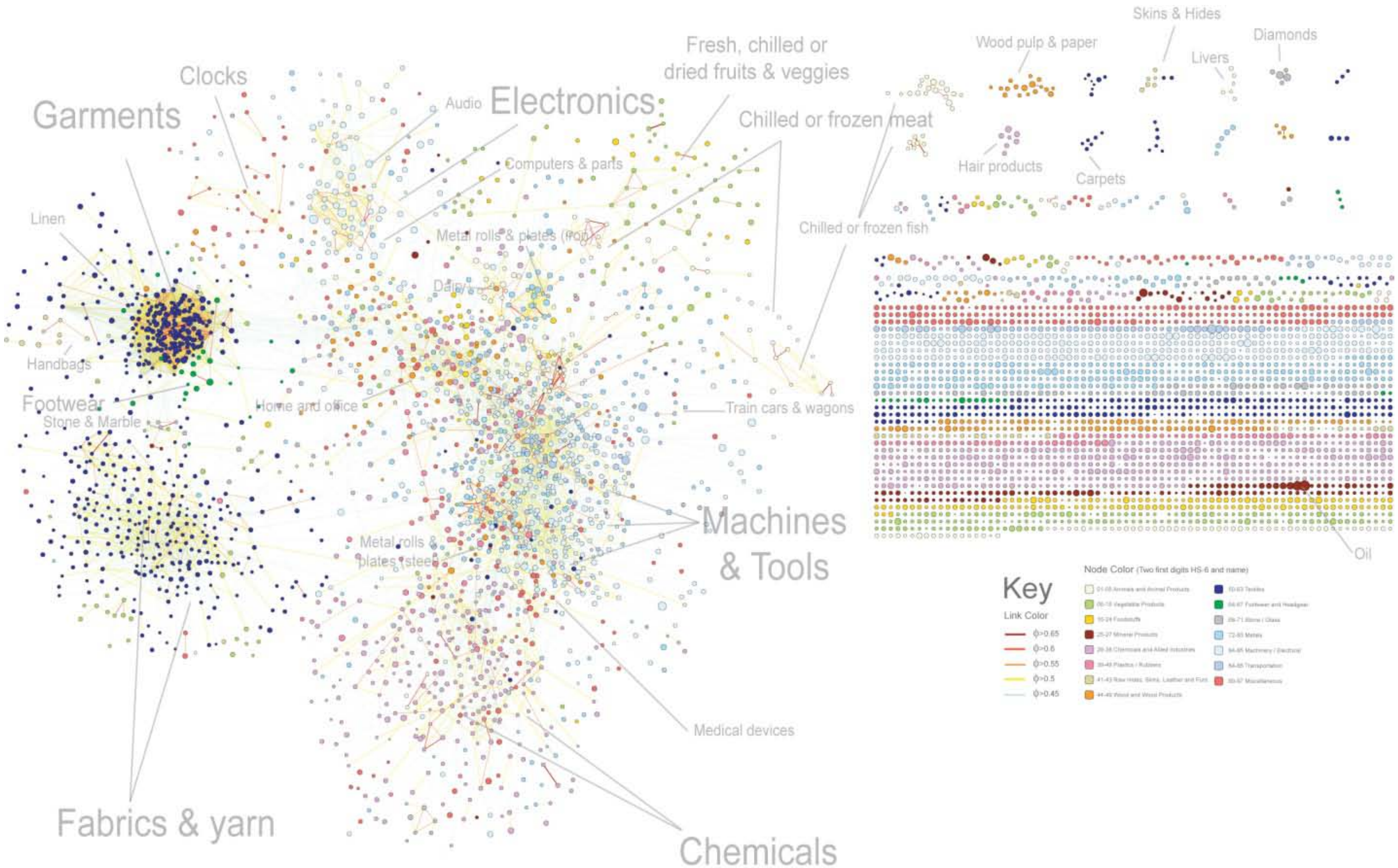

Figure 11 Visualization of the Product Space for the year 2005 (Data from Gaulier et Zignano 2009, Analysis and Visualization from Author's own calculations)

DATA

We use the HDI data compiled by Gray and Purser (2009) summarizing HDI and its components between 1970 and 2005 in five year intervals. When visualizing the time series for HDI and its components we limit ourselves to data only for countries for which data is available for all periods. As a result of this the number of countries included in visualizations for different indicators varies according to the availability of data. For the Partial Ordering Network we use data for 2007-2009, also compiled by Gray and Purser (2009).

THE HUMAN DEVELOPMENT INDEX

The Human Development Index (HDI) is a statistic used to describe countries well being that gives equal weight to (i) an education indicator, (ii) a life or health indicator, and (iii) an income indicator. Formally the HDI of country *c* at time *t* is defined as:

$$- \qquad - \qquad - \tag{1}$$

Where $ED_c(t)$, $LIFE_c(t)$ and $GDP_c(t)$ represent, respectively, an education indicator, which combines (2/3) of an adult literacy indicator and (1/3) of a primary enrollment indicator, a life indicator that is proportional to life expectancy and an income indicator which is proportional to the country's per capita Gross Domestic Product (GDP). All of these indicators are normalized such that their values lie between 0 and 1.

Starting in 2010 the HDI will be modified to include (i) School life expectancy of children of school entrance age and (ii) mean years of schooling of adults as its new educational indicator. The figures we present below, however, use the traditional definition of HDI described in equation (1) and use the data compiled by Gray and Purser referenced in the data section of this document.

(I)-(III) RANKINGS, TRENDS AND DISTRIBUTIONS

(i) Rankings

We begin exploring the evolution of nations' relative levels of human development by looking at the ranking of HDI and its components across time. Figure 8 to Figure 14 show the evolution of the ranking of countries according to HDI between 1970 and 2005. The figures were created by assigning colors to each country in 1970 according to a perfect gradient that goes from dark blue, for the highest ranked countries, to crimson for the lowest ranked ones. This coloring technique helps to quickly identify countries that have experienced substantial changes in their HDI ranking during this period. The smooth coloring of Figure 12 is evidence that, despite the substantial changes in the levels human development that took place during the last 30 years, the positions in the HDI ranking remain relatively stable. More notorious changes are observed in the rankings of individual indicators. For instance Figure 13 shows that Greece (GRC), Korea (KOR), Lybia (LBY), Indonesia (IDN) and Algeria (DZA), among other countries, moved substantially up the education indicator ranking, whereas Belize (BLZ), Vietnam (VNM) and Trinidad and Tobago (TTO) pop out as countries that fell down several spots in the education ranking.

Figure 14 shows the Life indicator ranking revealing that in this dimension the countries that moved up the ranking the most during this period were Korea (KOR), Oman (OMN), Vietnam (VNM), Indonesia (IDN) and Nepal (NPL) among others. For the income indicator, however, the countries that moved up the ranking the most were Equatorial Guinea (GNQ), Botswana (BWA), Thailand (THA) and China (CHN) (Figure 15). Movements in ranking, however, are not equivalent to movements in values. This is illustrated in Figure 16, which highlight the 15 countries that experienced the largest increase in HDI during this period, with numerical values shown in parenthesis. Figure 16 shows two things, the first one is that most of the countries that made substantial advances in Human Development during the observation period were all located in similar positions of the ranking. As a result, the ranking structure is such that some countries that made substantial progress during the period, such as Botswana (BWA) moved down the ranking. Moreover, the top performer of the period, Indonesia (IDN) moved up only a few positions, since all countries that were close to Indonesia in the ranking moved up substantially as well. The second thing that is clear from Figure 16 is that most of the countries that had largest increases in HDI are located either in East Asia or in North Africa and the Middle East. This geographic concentration of HDI increase is shown geographically in Figure 17.

Rankings such as the ones presented in Figure 12 to Figure 16 can also be used to visualize the relative position of regions as well as countries. Figure 18 to Figure 23 shows a visual exploration of the ranking of countries and the relative position of the regions they belong to by coloring each country according to its respective regions. Figure 18 shows the HDI ranking together with that of each of the three components of HDI highlighting the clear separation in the HDI rankings that exist between these different geographical regions. Figure 19 uses this figure to compare the evolution in HDI and its components experienced by Chile and Korea. The figure shows how Korea has surpassed Chile not only in the HDI ranking, but in each of the components as well, the largest difference in ranking between the two countries now being the difference in the education index and the smallest being that in the life component of the indicator.

Figure 20 to Figure 23 highlight the position in the HDI ranking and that of its components of countries located in Africa, Latin America and the Caribbean, Asia and Europe together with North America and Oceania. The patterns defined by each one of these regions are very distinct. Most African nations are consistently located in the bottom spots of the ranking, except for a few nations whose income depends on the extraction of oil such as Equatorial Guinea (GNQ), Lybia (LBY), Gabon (GAB); a exception being the Seychelles (SYC), a small archipelago in the Indian Ocean that has become a prosperous tourist destination.

The patterns defined by Latin America and Asia are, on the other hand, characterized by a high degree of heterogeneity. Countries from Latin America and the Caribbean occupy a wide range of the HDI ranking, with regional leaders in relatively high positions of the ranking, such as Barbados (BRB), Argentina (ARG), Chile (CHL) and Mexico (MEX) sharing the region with other countries like Haiti (HTI), Guatemala (GTM) and Nicaragua (NIC), who occupy some of the last positions in the ranking. This heterogeneity appears to be more pronounced in Asia, especially when it comes to the life and income indicators where some Asian countries can be found at the top of the ranking while others occupy the very bottom of it.

Figure 23 highlights Europe, North America and Oceania. While it is trivial to see from these figure the fact that these countries are concentrated in the top spots of the HDI ranking and that of its components it is surprising to see the relatively step decay on the Life ranking of countries such as those of the Former Soviet Union (FSU), Romania (ROM), Bulgaria (BGR) and Hungary (HUN). The decays in the ranking of these countries appear to have occurred between 1970 and 1995, after which these countries stop losing positions in the ranking or begun to recover a few spots in it.

Rankings like the ones presented in Figure 12 to Figure 23 are useful for doing some comparisons but are non-informative about HDI values. Moreover, rankings distort axis by stretching them around values where there is a large concentration of countries and compressing them around values where too few countries are located. This can create artificially large separations in the ranking between countries with similar HDI values while at the same time creating artificially small gaps between countries with extremely different values when no other countries are located between them.

In the next section we present a visual exploration of HDI values and also analyze changes in their distributions.

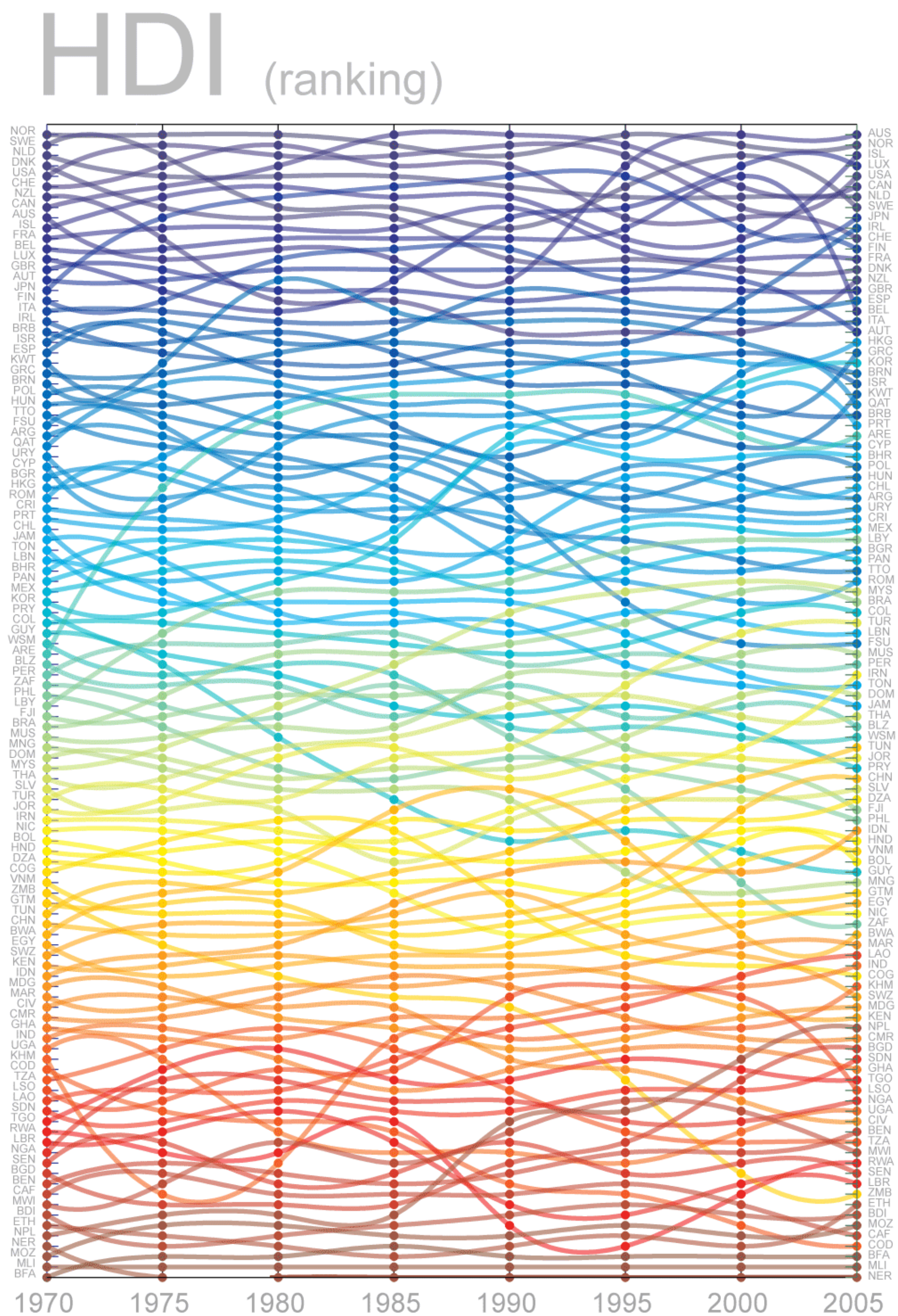

Figure 12 Evolution of the HDI ranking for the 1970 2005 period.

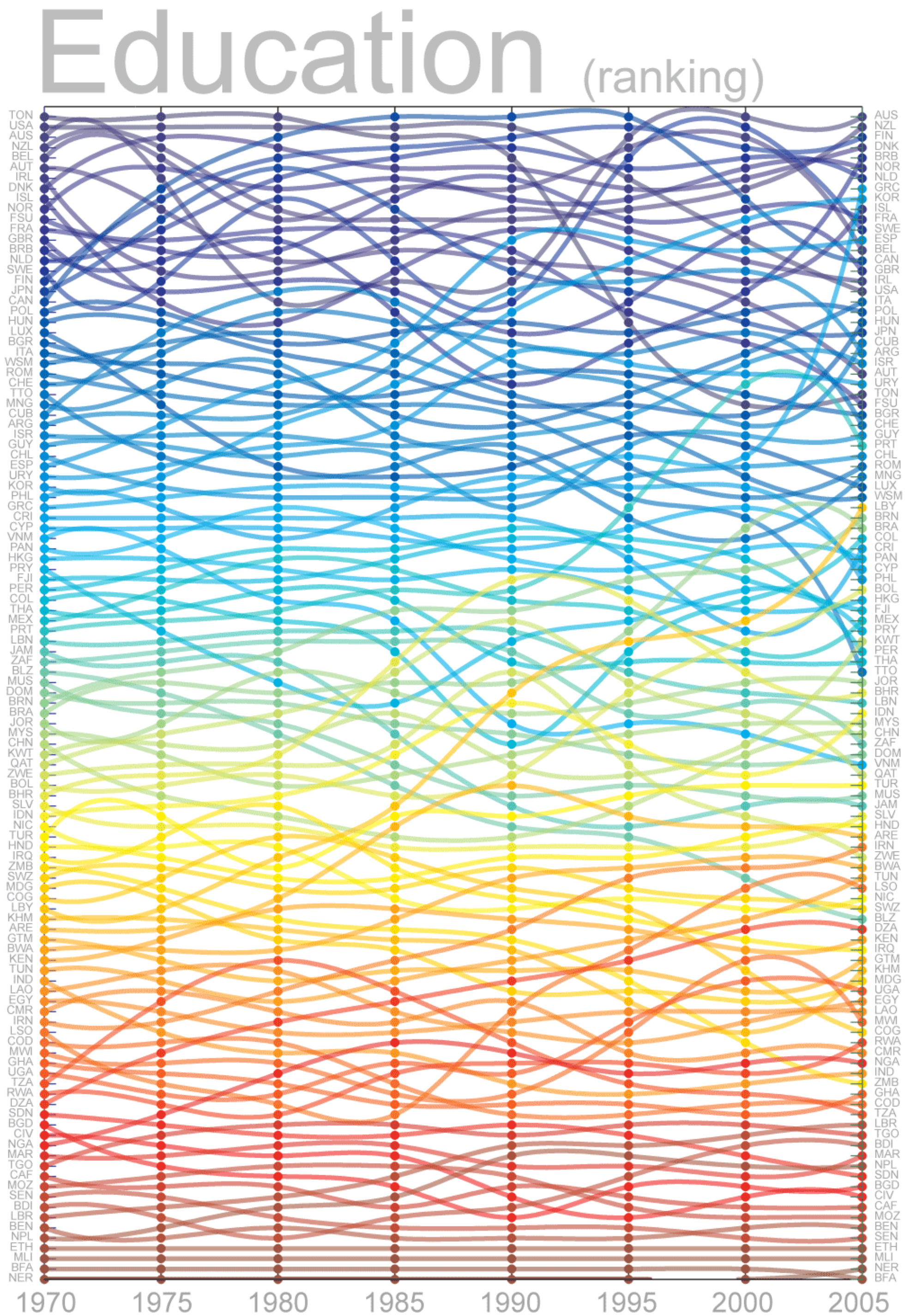

Figure 13 Evolution of the education ranking for the 1970-2005 period.

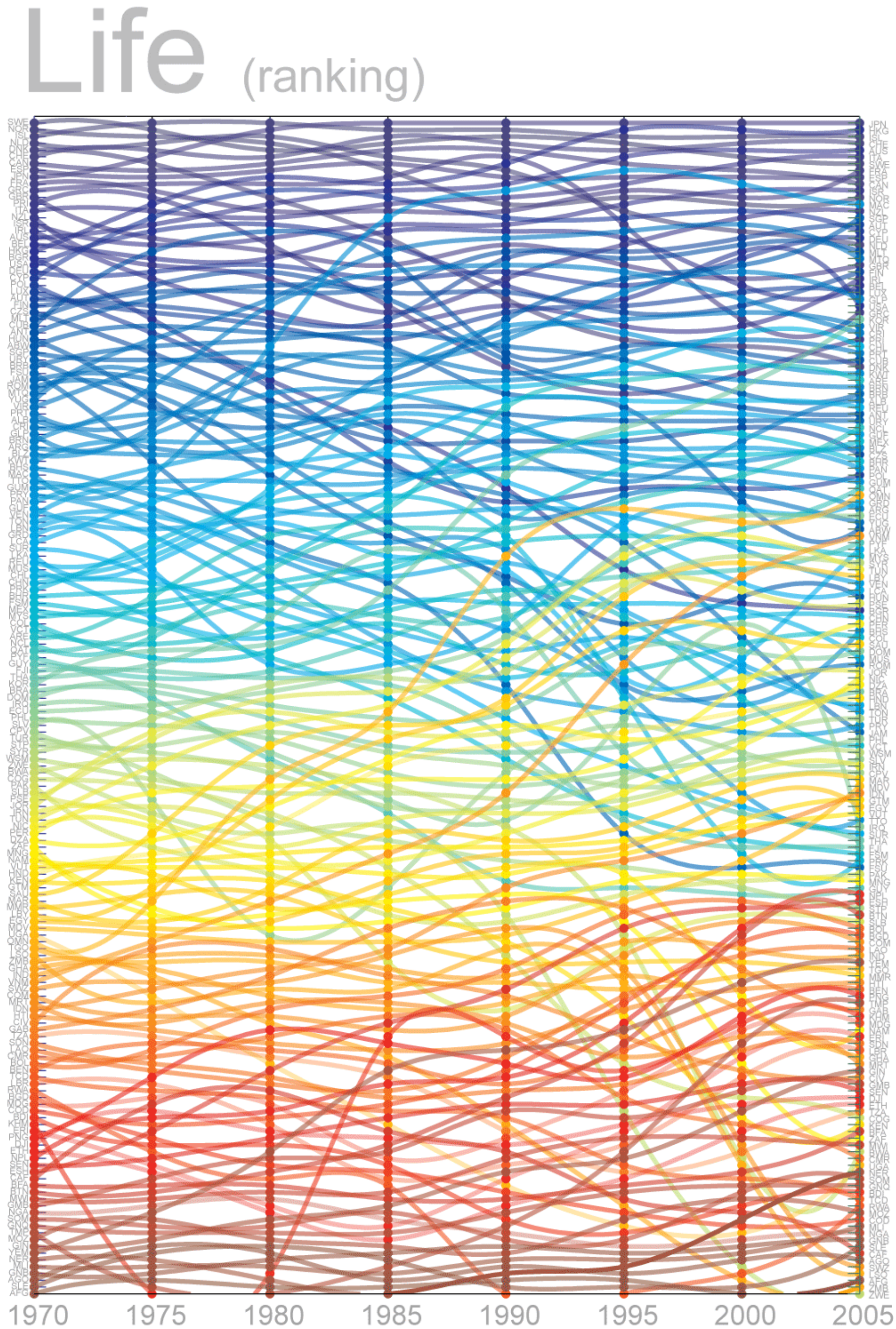

Figure 14 Evolution of the life ranking for the 1970 2005 period.

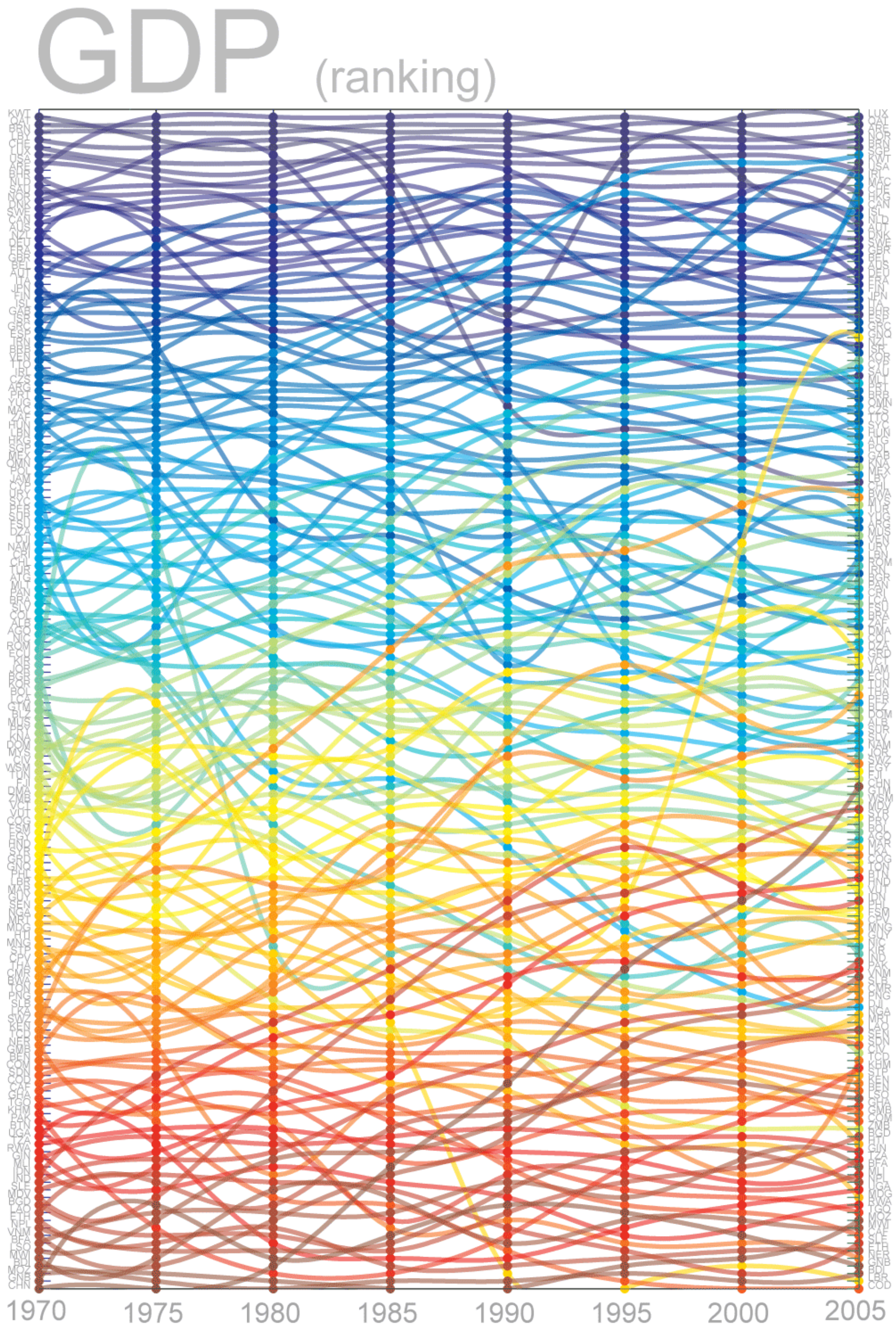

Figure 15 Evolution of the GDP ranking for the 1970 2005 period.

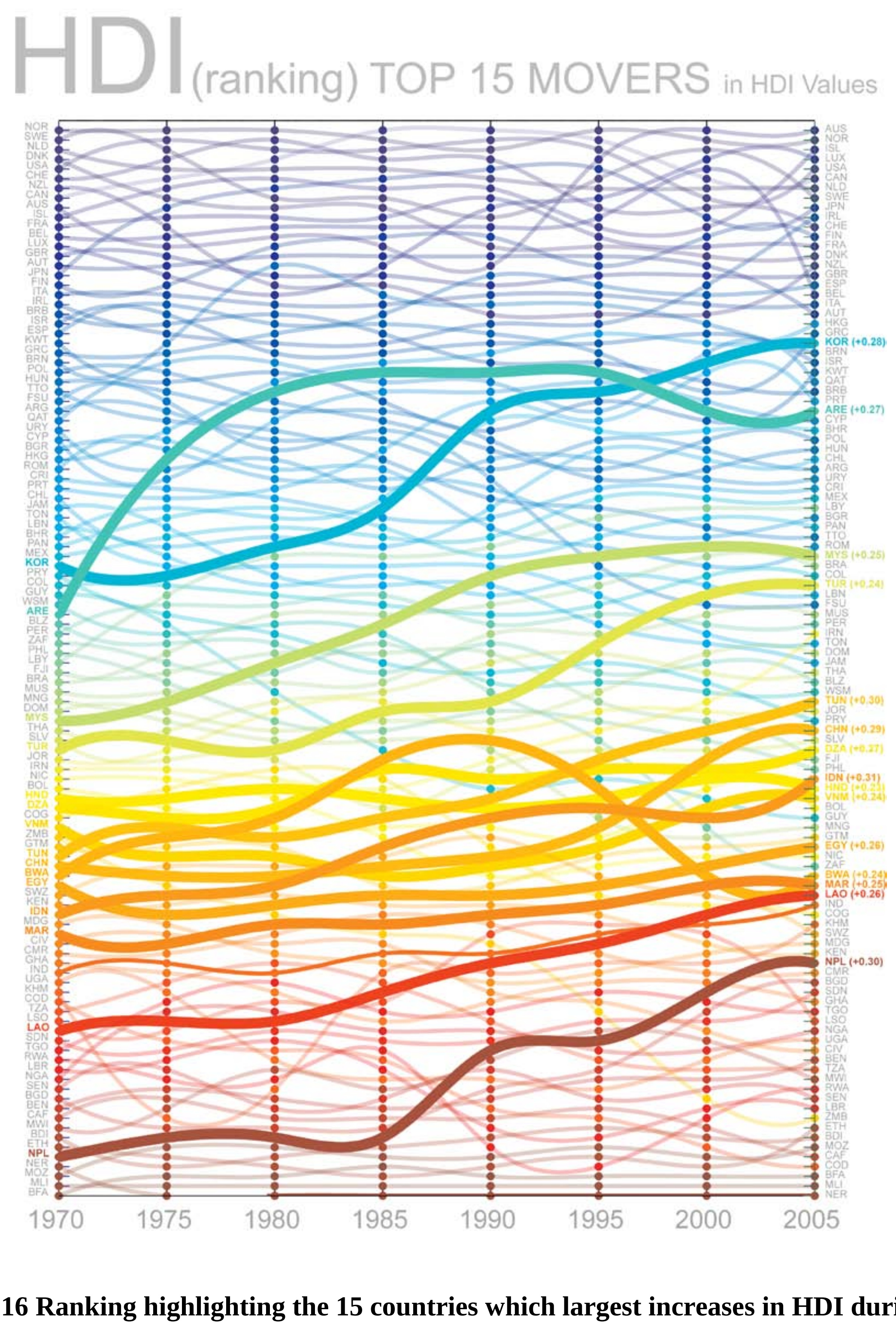

Figure 16 Ranking highlighting the 15 countries which largest increases in HDI during the period between 1970 and 2005.

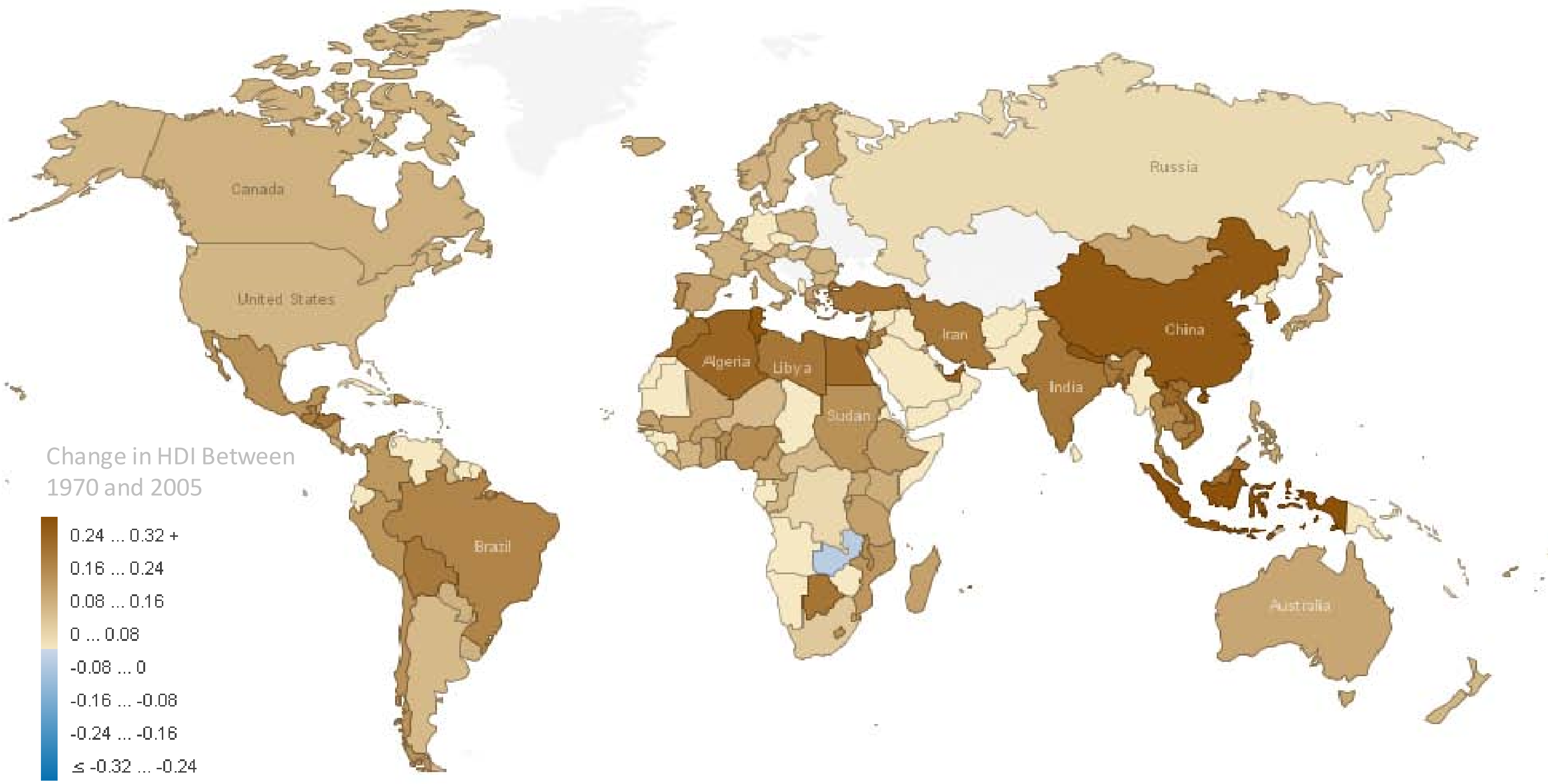

Figure 17 Figure illustrating countries with the largest changes in HDI between 1970 and 2005

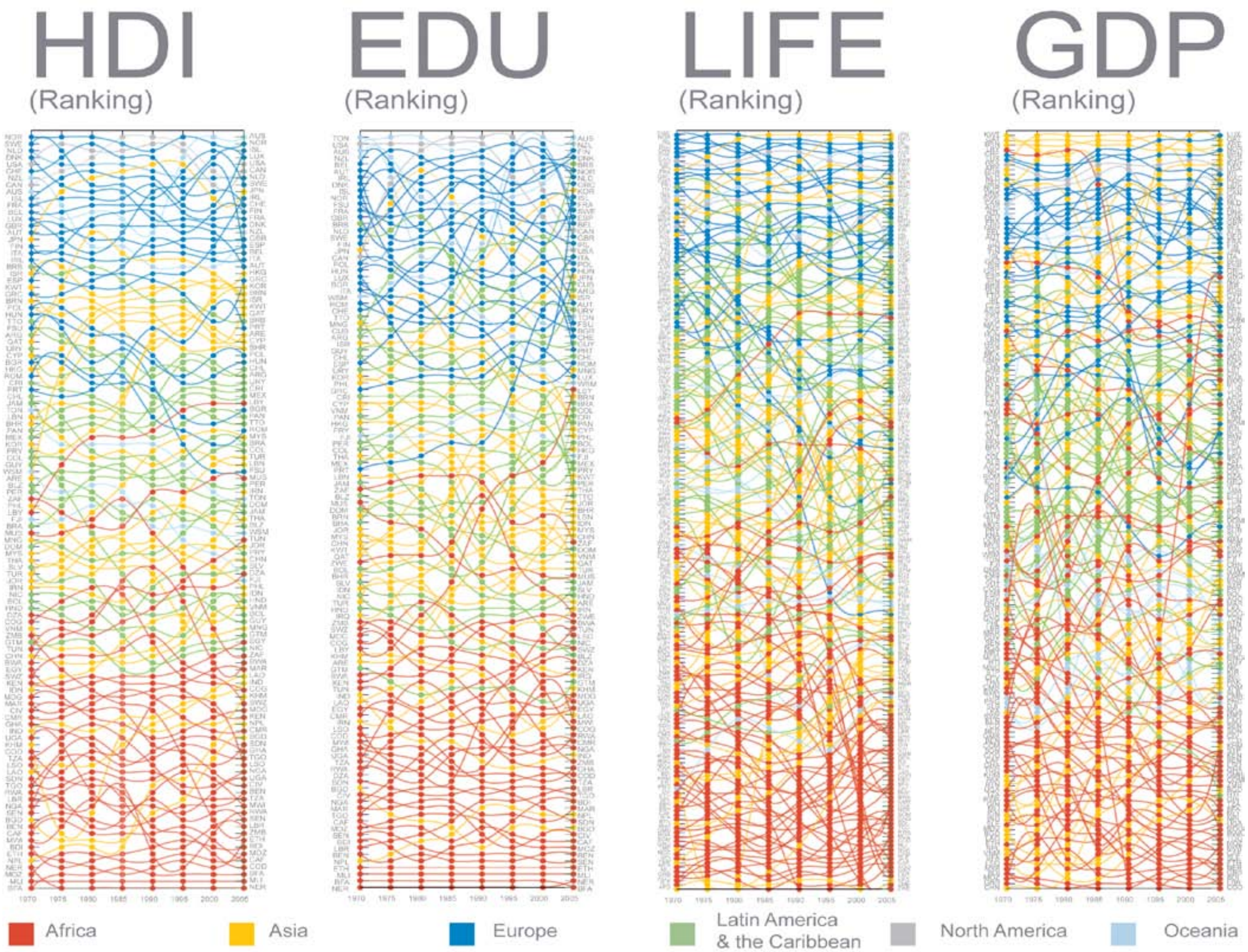

Figure 18 Evolution of the ranking of countries according to HDI and each of its three components with countries color coded by regions for the 1970-2005 period.

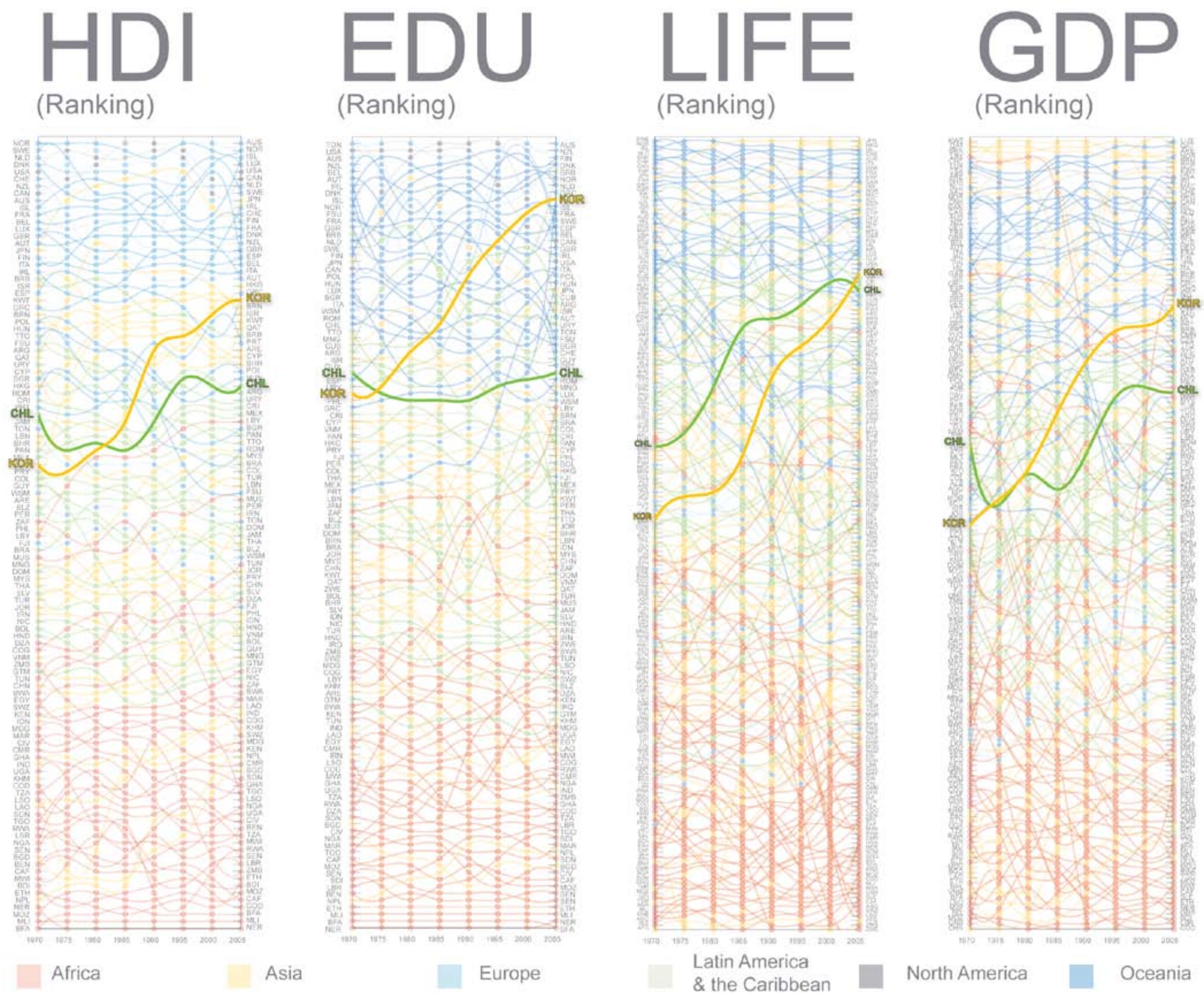

Figure 19 Comparison between the evolution of ranking of Chile and Korea in HDI and its components.

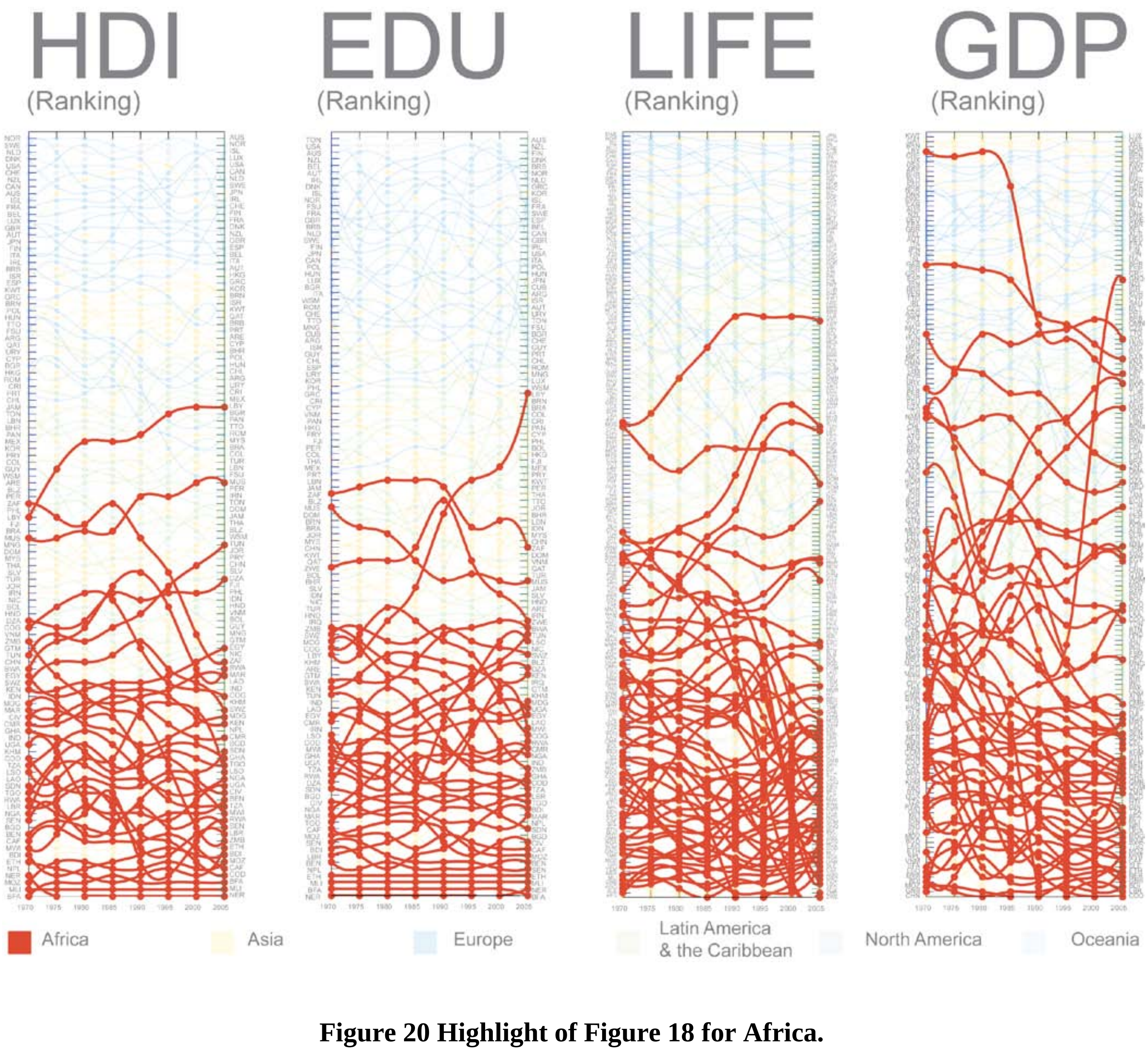

Figure 20 Highlight of Figure 18 for Africa.

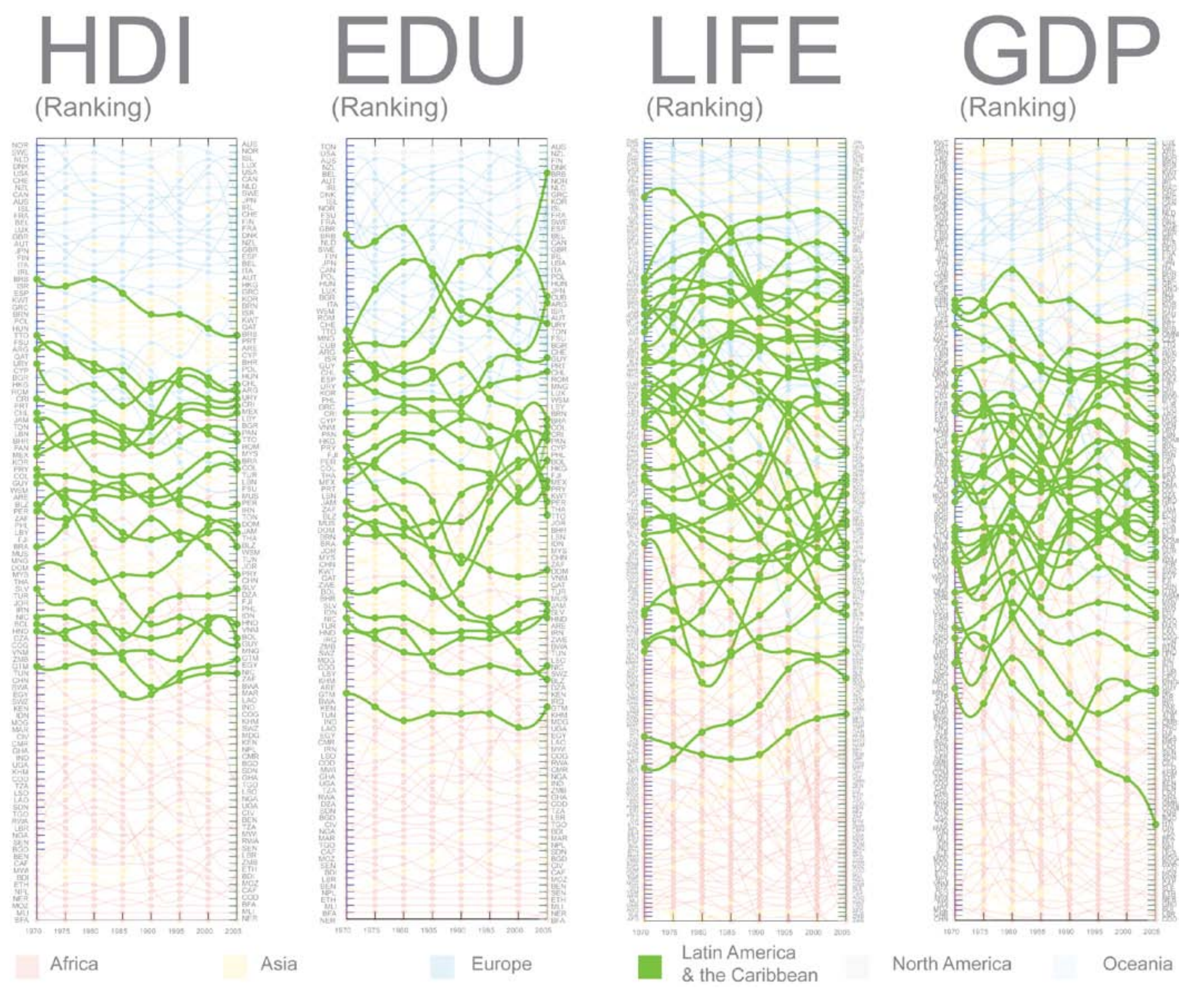

Figure 21 Highlight of Figure 18 for Latin America and the Caribbean.

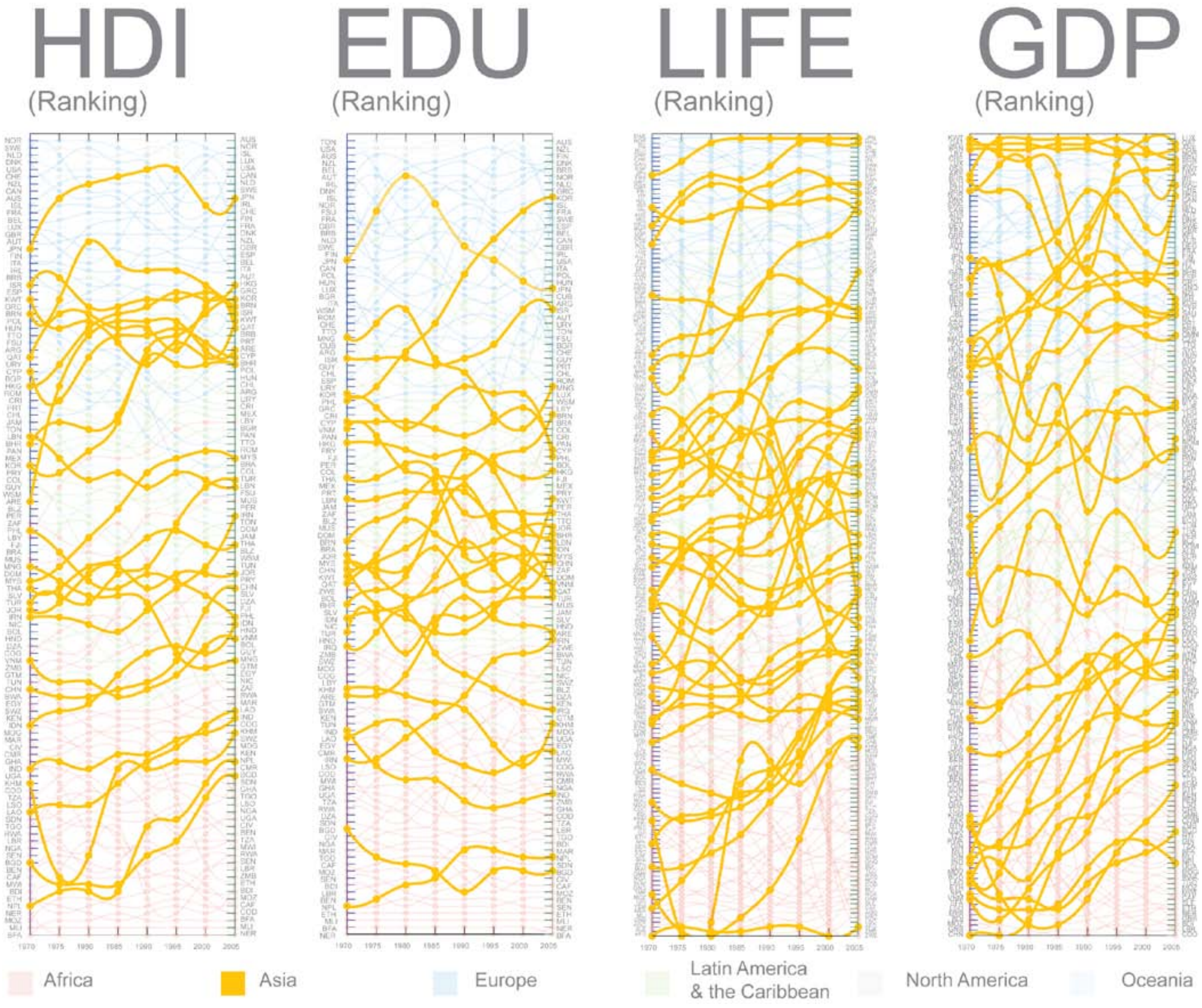

Figure 22 Highlight of Figure 18 for Asia.

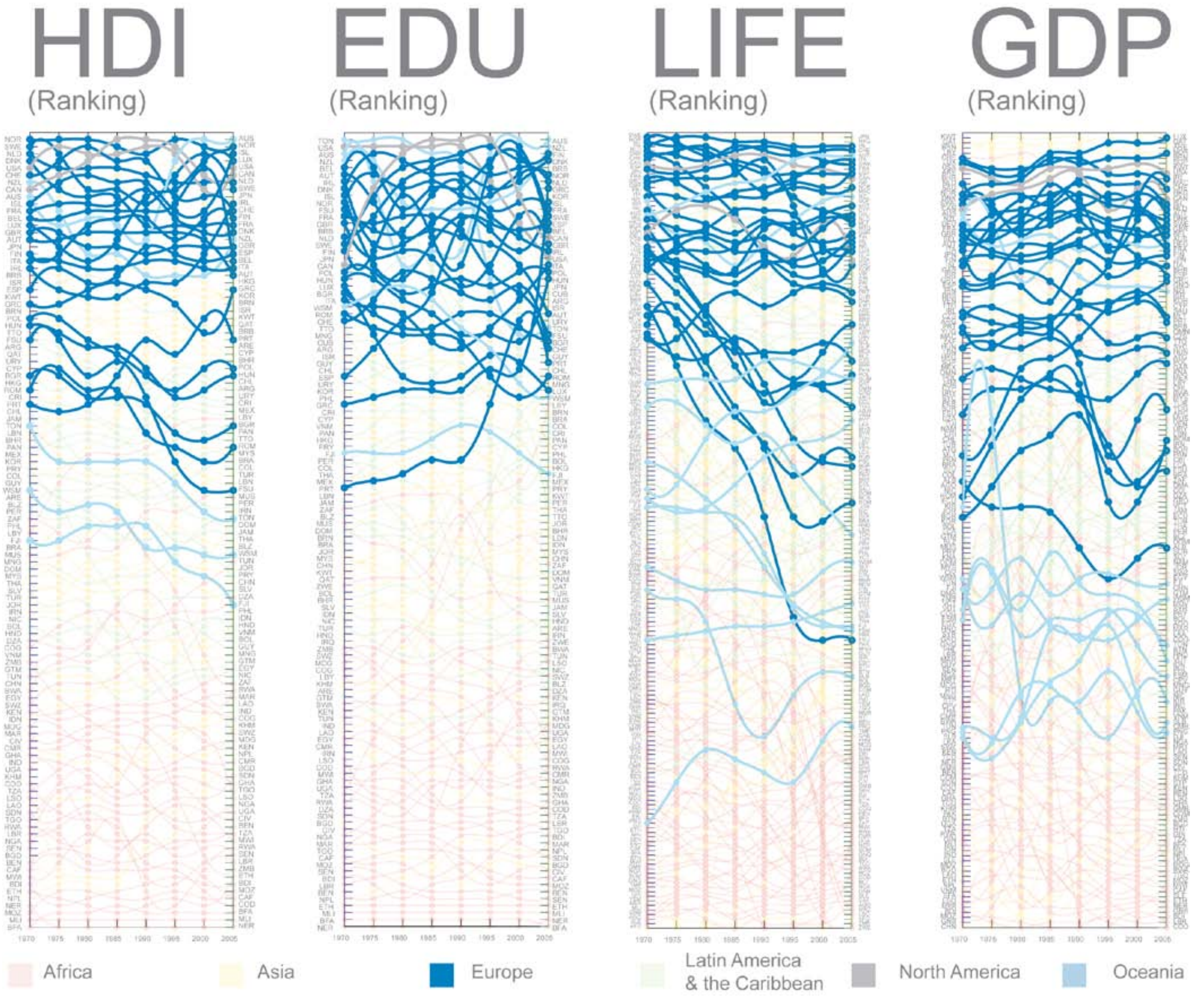

Figure 23 Highlight of Figure 18 for Europe, North America and Oceania.

(ii)-(iii) Values and Distributions

Figure 24 to Figure 28 show the evolution of HDI values for all countries for which we have data available between 1970 and 2005. The upward slopping curves of these figures signal a substantial progress in human development, as captured by HDI data, and its components, during these 25 years.

These figures (Figure 24 - Figure 27) also show the rankings of countries in the axis, evidencing a high concentration of nations at high HDI values. This is especially true for the education and life indicators, suggesting a need to promote these measurements or complement them with additional ones if they are to continue helping us distinguish between the level of development of countries and helping identify specific problems affecting countries that have "outgrown" the current Human Development Index.

The relative concentration of countries at some HDI values can be appreciated better by looking at figures showing the bivariate probability of finding a country with a certain HDI value for a given year. Figure 29 shows the evolution of the distribution of HDI values, and its components, between 1970 and 2005, for all countries. This figure illustrates some interesting stylized facts regarding the distribution of HDI values. First, they show relatively high concentration of countries for values close to maximum, especially for HDI and the education indicator, for which more than half of the countries are concentrated at values of the indicator of 0.8 or higher.

We can also get a deeper look at each region by recreating these distributions using data for countries in a single region. Figure 29 does this exercise for the HDI density distribution highlighting the different patterns characterizing Europe, Latin America and the Caribbean, Asia and Africa. Once again this figure illustrates the large regional differences in HDI values characterizing countries of different regions. Fig 29 shows, for instance, that HDI almost does not capture any differences between European countries, all of which are concentrated at extremely high HDI value (above 0.9), while at the same time it reveals a large degree of heterogeneity in HDI among Asian states.

All in all, the visual exploration of HDI values and their distribution shows that (i) there was a substantial increase in development as captured by HDI between 1970 and 2005 (ii) this increase has resulted in the saturation of some dimensions, with many countries located in a narrow range of values close to the maximum. This is specially the case for HDI and the Education indicator, but there is also a large concentration of values in the life dimension as well. These suggest a need to extend or complement these indicators with other measures, a step that has been taken regarding the education indicator. Additional measures for extending the life dimension could include the prevalence of diseases that have reached epidemic proportions, such as AIDS and obesity, as well as indicators of sanitary conditions, such as the availability of drinking water, for instance. The extension of the indicators would increase the ability of HDI and its components to differentiate between countries and point to more specific problems. (iii) There are important differences in the variations of level of development, and the structure of this variation, for countries in the same region. Europe, for instance, appears in the distribution of HDI values as a coherent mass of countries that has progressed steadily during the observation period towards the highest values of the ranking. Asia and Latin America, on the other hand, exhibit a wide range of variation which is characterized by bands of countries. Asia in particular shows three well defined bands in Figure 30 that indicate that the Asian continent is composed by a set of developed countries, including Japan (JPN), Korea (KOR), Israel (ISR) and Kuwait (KWT), which have made a comparable progress to that of Europe and North America, followed by a second set of countries that have adequate levels in HDI and its components, such as Thailand (THA), Malaysia (MYS), China (CHN) and Turkey (TUR). Finally there is a third group of countries that has grown at a similar rate but still lags considerably behind the rest, including examples such as India (IND) and Laos (LAO).

Here we conclude our use of traditional graphical statistical methods. In the next sections we explore the use of a symbolic and network representation for the Human Development Index. In the next section we use a tree as a visual metaphor to illustrate HDI together with its components. In the section after that we introduce the concept of a partial ordering network to graphically illustrate the patterns defined by connecting countries at comparable levels of development.

BOX 1:

GRADUATION FROM HDI?

The remarkable progress in HDI and its components exhibited by several countries during the last couple of decades has translated into a large concentration of countries in a narrow range of relatively high HDI values. This concentration limits the ability of HDI to point into particular problems and limits the ability of the indicator to distinguish between countries.

As progress in human development is made it is natural to think that countries will “graduate” from the indicator. For instance, it is hard to believe that measures of life expectancy, school enrollment or adult literacy could help identify or signal any problems for many of the countries in the dataset. The need for “graduating” many nations from the education indicator is well exemplified, for instance, by the fact that in 2005 the education indicator was the largest contributor to HDI for 75 (67.6%) out of the 111 countries for which full HDI data was available. In other words, the education indicator, composed of Adult Literacy and Primary Enrollment, is a dimension in which it is relatively easy for all countries to obtain a large score and does not help discriminate between different nations.

It can be useful to think of HDI using an analogy in which HDI is a test and countries are students. Currently, the test consist of three questions, from which the education question is one that most students get right and therefore does not contribute much to the discriminatory power of the test. This suggests a need for the subcomponents of the education question to be upgraded, to "make the question more challenging" in order to separate the best students, from the good ones and the good ones from the more mediocre ones.

The graduation and extension of the Life indicator also seems appropriate at this point, albeit for a difference set of reasons. While most countries are not concentrated in extremely large values of the Life indicator, most countries are still concentrated in a narrow range of the Life indicator values. Indeed, in 2005 the standard deviation of the Life indicator for was 0.182, only slight larger than the standard deviation of the education indicator, which was equal to 0.177 for that same year. For comparison, the income indicator has a standard deviation of 0.224. The narrow range of values taken by countries in the life indicator suggests a need to extend or improve this dimension to include sanitary and health related issues that are not captured by the current incarnation of the Health Indicator.

The graduation and extension of the Human Development Index does not need to be traumatic. *Au contraire*, it can be done gradually and respecting HDI's current tripartite structure (i.e.

conserving the branches of education, life and income). This would keep HDI simple for explanatory purposes, at an aggregate level, as it would conserve its structure as an average of education, life and income indicators, and would move the sophistication of new dimensions into higher levels of the indicator, where the definitions are already more complex (as they include the rescaling procedures).

The inclusion of more dimensions into each of the three branches of the HDI is an exercise that should be performed carefully by a panel of experts that would consider data availability, as well as cultural differences and other considerations. Yet, suggestion for more modern incarnations of the education indicators could include, for instance, higher forms of education, such as high school graduation and university enrollment as well as availability and access to information technologies, including the number of individuals with access to internet and mobile phone, for instance. The life indicator could be improved using, public health dimensions such as access to clean water, or the state of epidemics such as that of AIDS, malaria and obesity.

All in all, a strategy for gradually extending an improving the components of HDI is necessary. As it is the indicator is not adaptable and runs the danger of becoming obsolete. A famous quote of Lewis Carrol, which evolutionary biologist love to mention is that "..it takes all the running you can do, to keep in the same place". The current statistics of the Human Development Index suggest that it may well be time to listen to the "red queen's" advice.

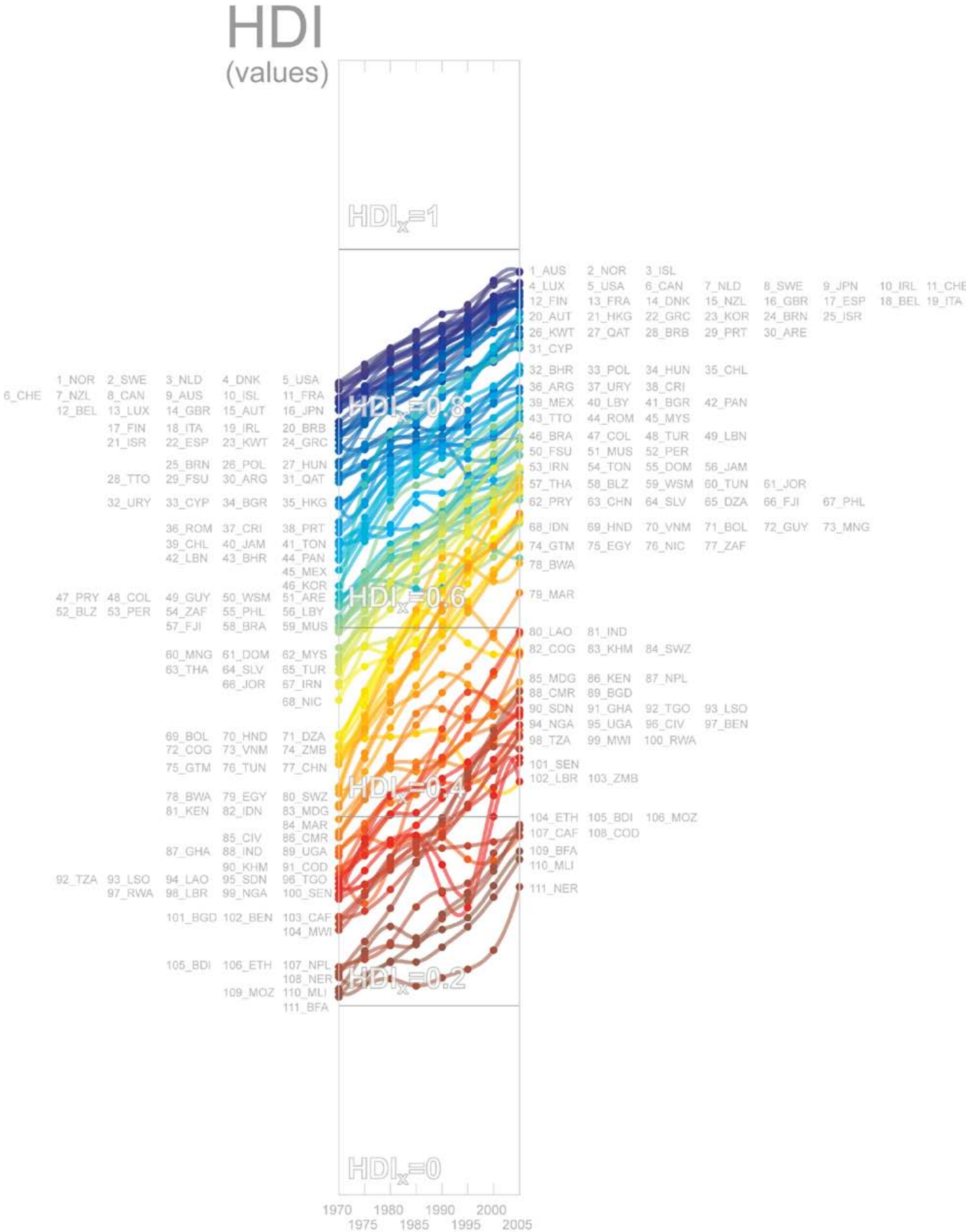

Figure 24 Evolution of HDI Values between 1970 and 2005

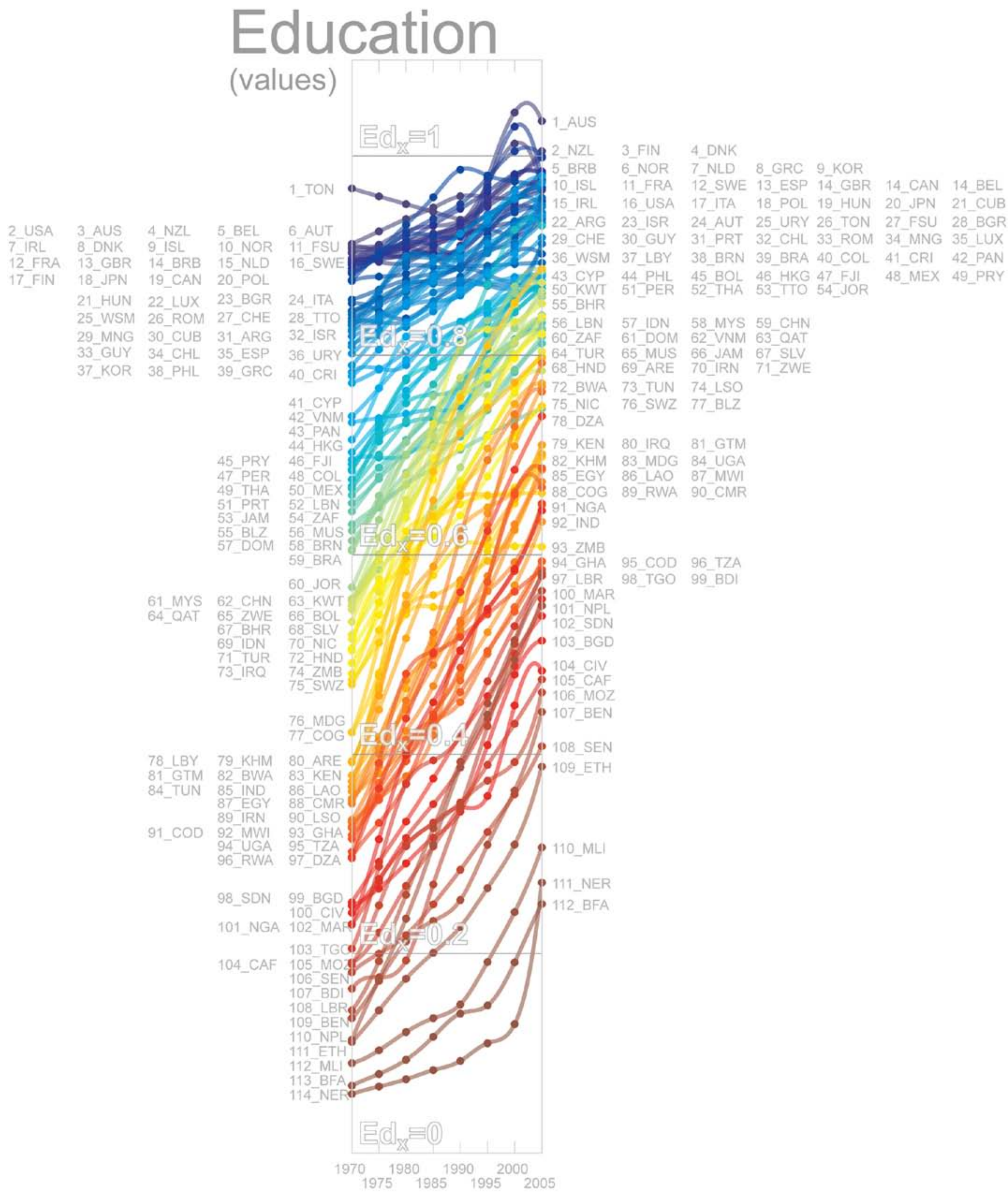

Figure 25 Evolution of the education component of HDI between 1970 and 2005

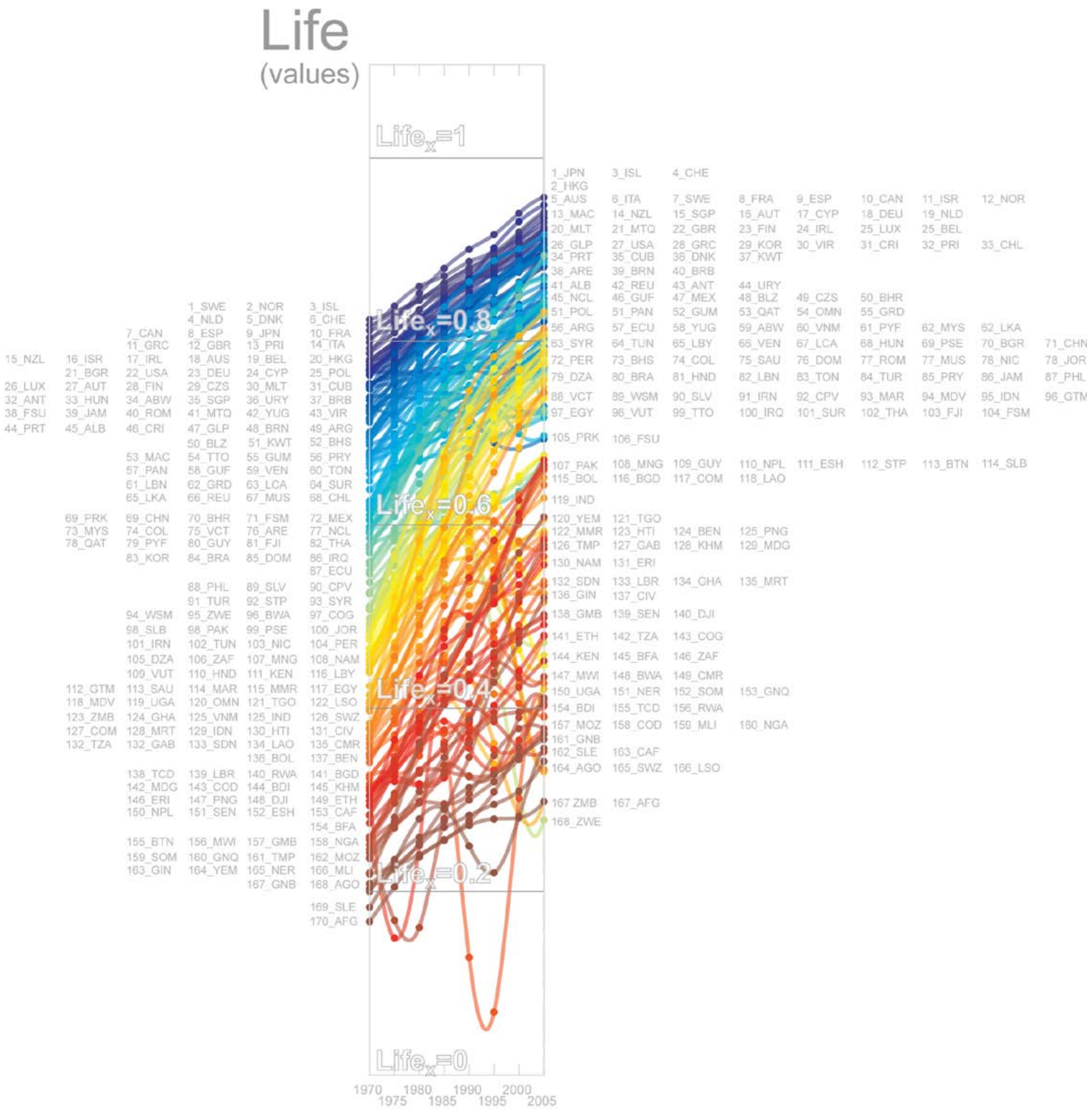

Figure 26 Evolution of the life component of HDI between 1970 and 2005

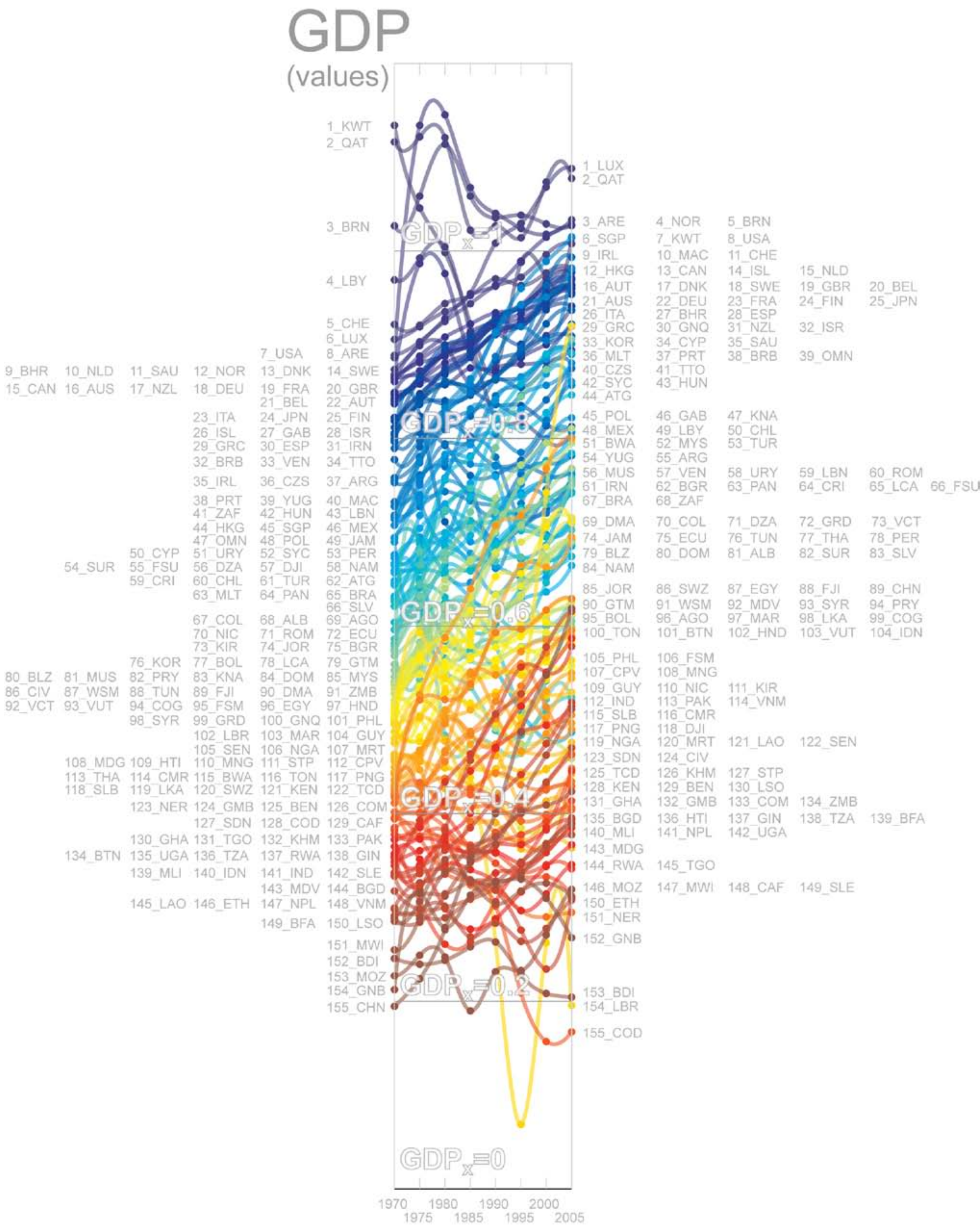

Figure 27 Evolution of the GDP component of HDI between 1970 and 2005

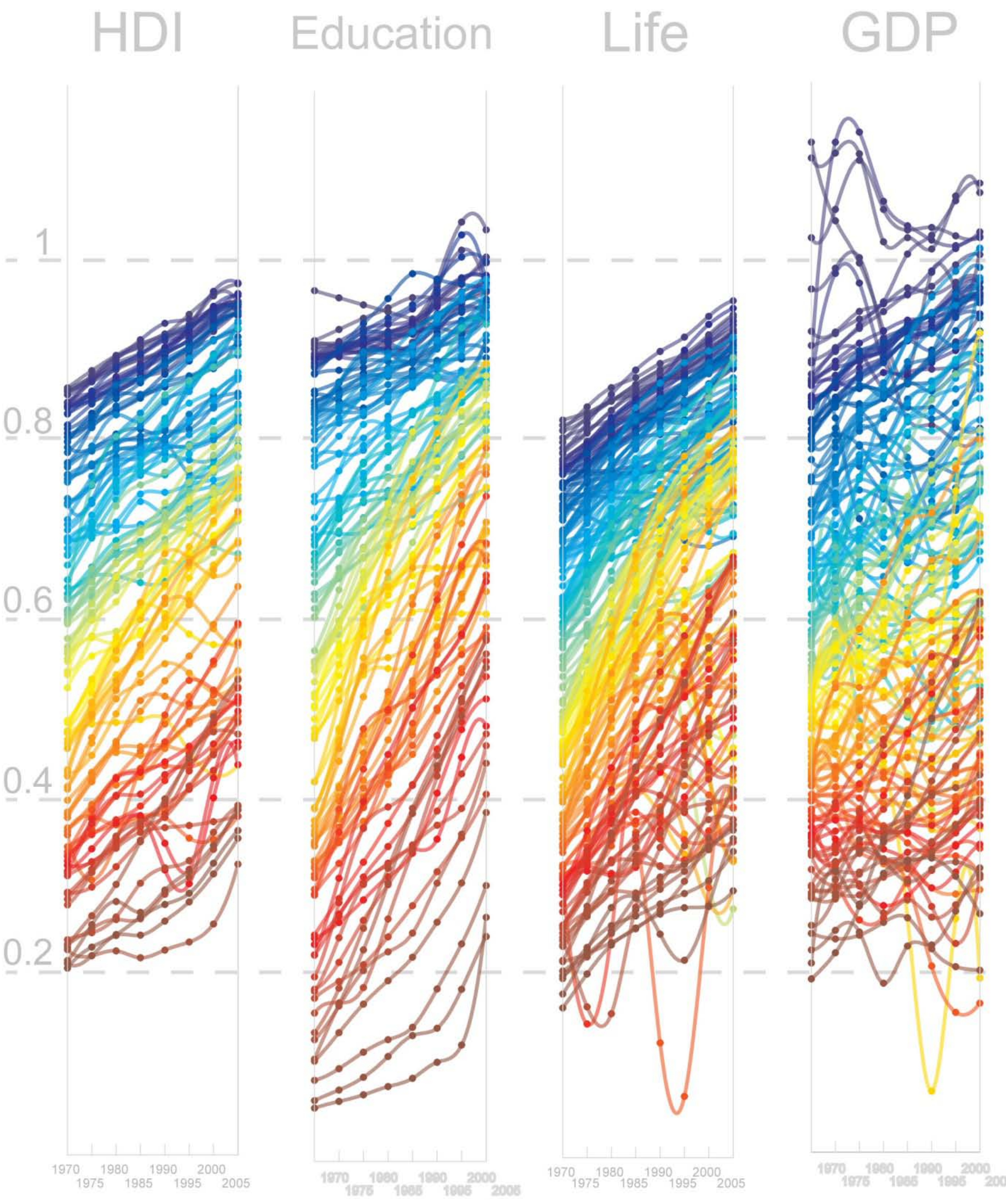

Figure 28 Composite of Figure 24 to Figure 27

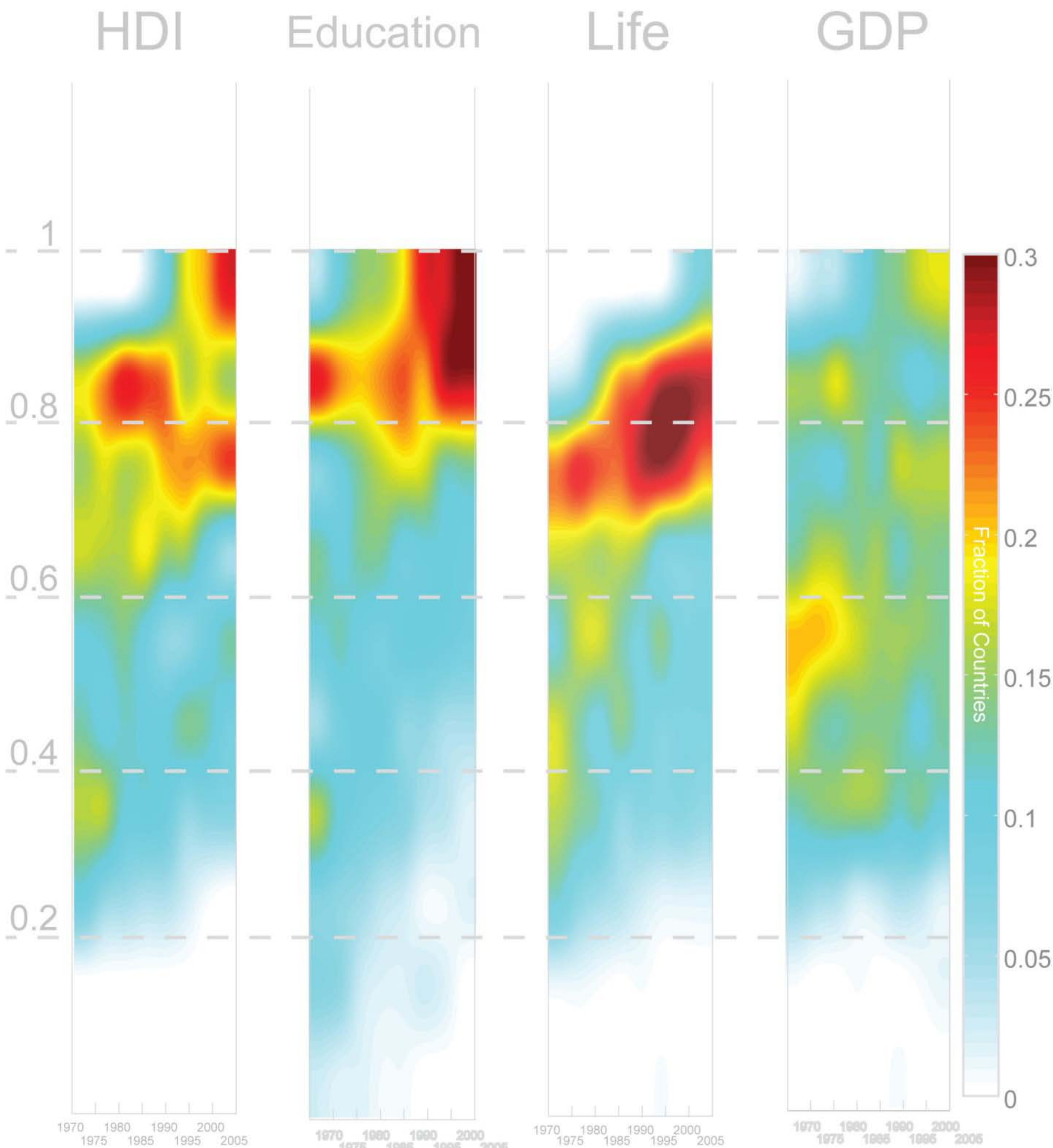

Figure 29 Evolution of the distribution of values of the Human Development Index and its components between 1970 and 2005

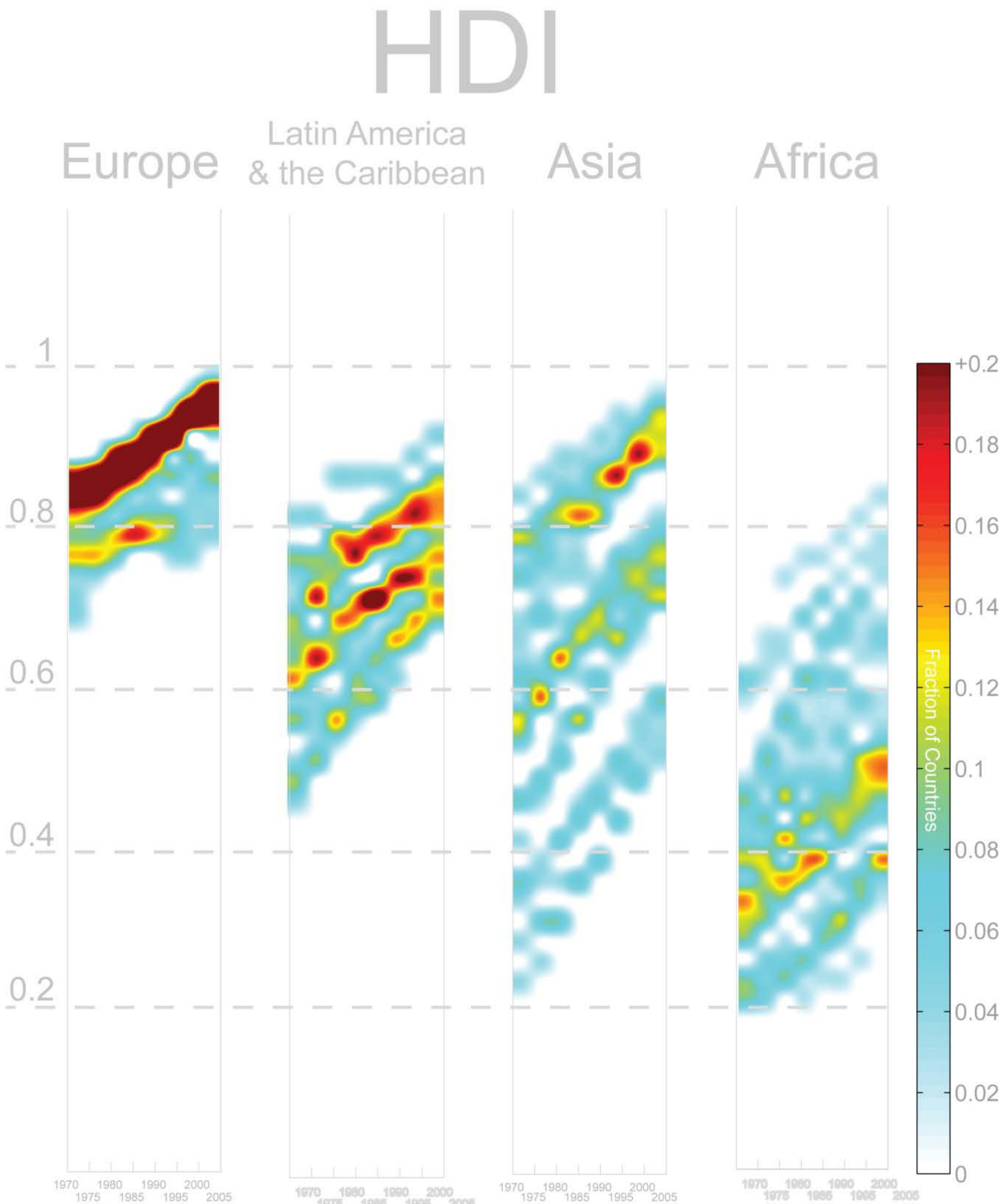

Figure 30 Evolution of the distribution of values of the Human Development Index between 1970 and 2005 for Europe, Latin America and the Caribbean, Asia and Africa.

(IV) THE DEVELOPMENT TREE

In this section we present a symbolic representation of the Human Development Index which we call The Development Tree.

From its definition, the Human Development Index is a “nested” indicator. This means that the value of HDI is the average of three other values, one if which is a weighted average of two other values. Because of this structure we can represent the formula used to determine HDI values using the diagram presented in Figure 31.

$$\text{HUMAN DEVELOPMENT INDEX} = \begin{cases} \text{(1/3) EDUCATION INDICATOR} \begin{cases} \text{(2/3) ADULT LITERACY INDICATOR} \\ \text{(1/3) PRIMARY ENROLLMENT INDICATOR} \end{cases} \\ \text{(1/3) LIFE INDICATOR} \\ \text{(1/3) INCOME INDICATOR} \end{cases}$$

Figure 31 Graphical representation of the formula used to estimate HDI.

The diagram presented in Figure 31 makes it clear that HDI can be represented as a tree with (i) a trunk given by the Human Development Index value (ii) three branches, given by each of the three HDI components, and (iii) two secondary branches representing adult literacy and primary enrollment.

Figure 32 shows the first of these two designs, which we call the Spline Design. The Spline Design was constructed according to the following design rules:

<u>Development Tree Rules (Spline Design)</u>

1. The height of the trunk is linearly proportional to the HDI value
2. The order in which the branches come out from the tree indicates the relative contribution of each component to HDI. The bottom branch being the smallest relative contribution and the top branch being the highest. For instance, in this example the life indicator is the one that makes the smallest contribution to the overall HDI while the GDP indicator is the one that contributes the most.
3. The point at which the branches begin (change of color) is proportional to its contribution to the total HDI.
4. The length of the branch (after the curve) is propotional to the actual value of the component indicator.
5. The color of the trunk is a weighted average of the color of the HDI components.

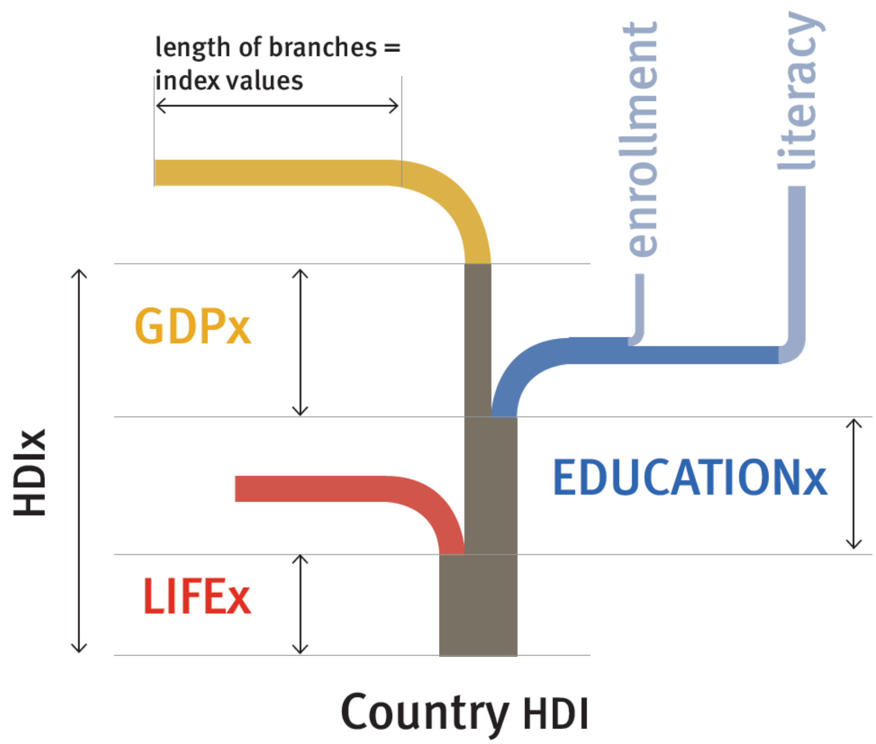

Figure 32 The Development Tree (Spline Design)

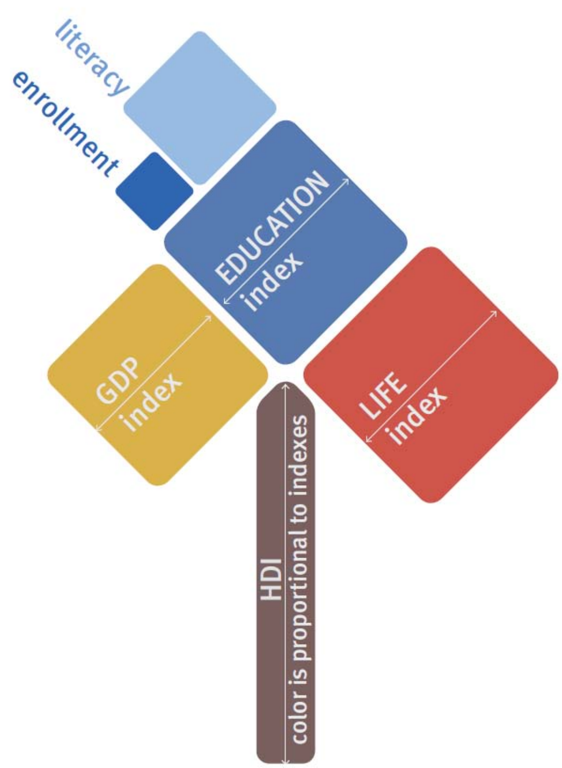

Figure 33 The Development Tree (Diamond Design)

Figure 33 shows our second design of the Development Tree, which we call the Diamond Design. The Diamond Design was constructed according to the following set of rules:

Development Tree Rules (Diamond Design)

1. The height of the trunk is linearly proportional to the HDI value
2. The order in which the branches (boxes) come out from the tree indicates the relative contribution of each component to HDI. The leftmost branch being the smallest relative contribution to the value of HDI and the rightmost box representing the highest relative contribution.
3. The side of the box is proportional to the actual value of the component indicator.
4. The color of the trunk is a weighted average of the color of the HDI components.

Unlike the figures shown in previous sections the development tree allows the visualization of HDI together with its components in a single image, providing a picture of the HDI value that can be immediately tied to the contribution to the total HDI value of each of the indicator's components.

It is also a plausible hypothesis that the iconographic character of the development tree will create a stronger emotional response in those experiencing the data than numerical or line graph representations. This makes the development tree a more suitable tool for teaching and illustrating the values of human development in public places and to non-technical audiences.

Because of both, its high information content and its potential ability to cause more emotional responses, the development tree can be used as an effective tool to create (i) graphical narratives of countries histories and (ii) comparative graphics for sets of countries.

Figure 34 uses the Diamond Development Tree Design to graphically illustrate the story of Rwanda. The figure clearly tells a story of a country that in 1970 had a small value of HDI which was evenly distributed among its three components, with education making the smallest contribution to HDI and income the largest. The image shows that by 1975 education had become the largest contributor to Rwanda's HDI which remained small and balanced. By 1985 the illustration shows that Rwanda had made progressed in HDI, most of which was coming from its improvement in Adult Literacy and Primary Enrollment. The 1990 and 1995 graphic, however, show evidence of the terrible genocide that affected Rwanda which is represented by the almost disappearance of the red box representing the life indicator. The image finishes by showing that by 2005 Rwanda had gotten back into a growing path of development. Figure 35 tells Rwanda story using the Spline Development Tree Design.

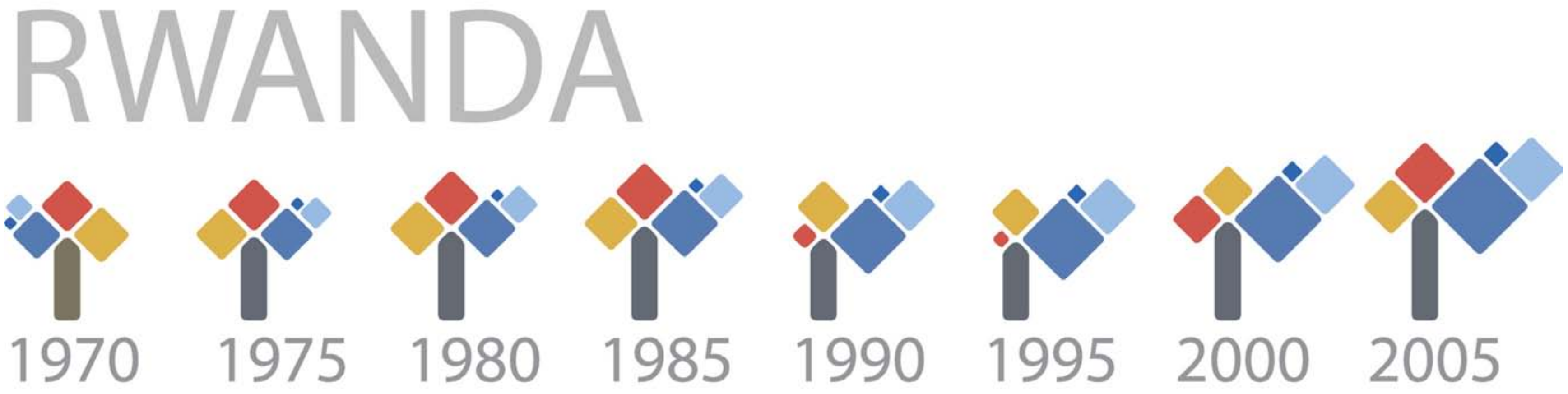

Figure 34 The evolution of the Human Development index and its components for Rwanda between 1970 and 2005 represented using the Diamond Development Tree Design

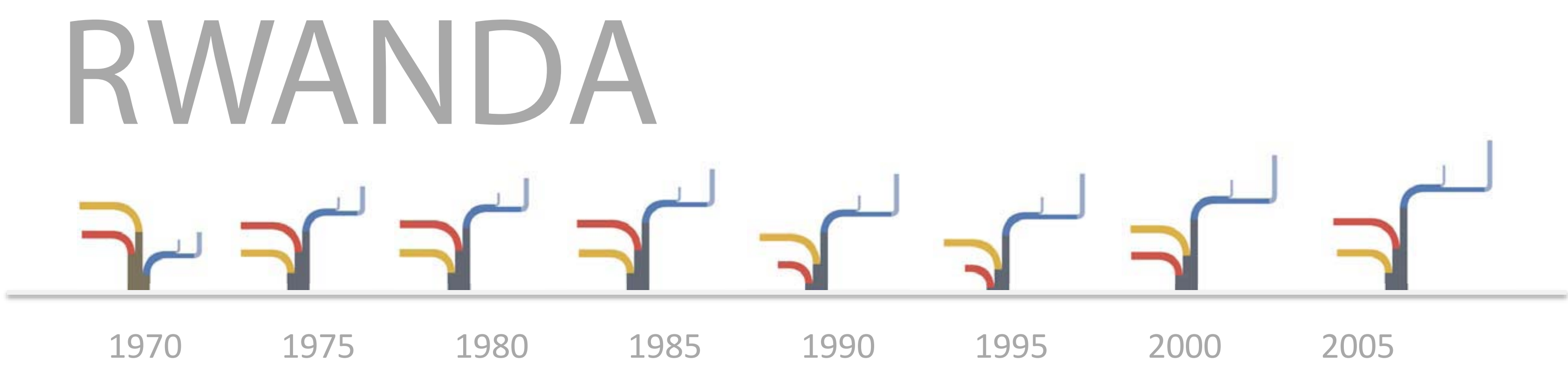

Figure 35 The evolution of the Human Development index and its components for Rwanda between 1970 and 2005 represented using the Spline Development Tree Design

The development tree can also be used to create graphical narratives for entire regions, or collections of countries. Figure 36 shows a poster comparing the Development Trees of 35 African countries, between 1970 and 2005. These visualization is excellent to quickly find out the countries that made more progress during this period, a set that includes Libya, Botswana, Egypt and Tunisia, among others. It also allows to quickly spot imbalances in HDI components, such as that exhibited by Botswana, evidencing the AIDS epidemic affecting that country, or the proportionally low income of the Democratic Republic of the Congo. Figure 37 shows the same as Figure 36 but using the Spline Development Tree Design. From this comparison, it is clear that the Diamond Tree is a better way to convey this large amount of information. We note that both of this images graphically summarize 432 different numbers, a quantity that is equivalent to a full page table. Yet, they help present this information in a way that makes it more amenable for comparisons and a wide audience.

Inequalities in the Human Development Index and its components can also be visualized using the Development tree. Figure 38 shows three examples of how opacity can be used to encode inequality in the total value of HDI or differentially in each one of its components. The use of opacity to encode inequality is based on the metaphor that more equalitarian distributions of HDI and its components are more "solid" than less equalitarian distributions.

COMBINING RANKING, VALUES AND SYMBOLIC REPRESENTATIONS

Finally, we present two visual explorations that combine information on HDI Ranking, HDI values and the Development Tree. The first of these representations, shown in Figure 39, compare the HDI values representing African countries in 1970 and 2005. The figure shows that most countries, except Zambia, have gone up in values, the value of this information visualization design, however, is the fact that it also shows the relative ranking of countries and uses arcs as a visual metaphor for the magnitude of the changes in HDI values. Green arcs show countries that went up in HDI value and ranking, whereas purple arcs show countries that went up in HDI value but lost positions in the ranking. As all arcs have an identical circular curvature the length of the arc is directly proportional to the change in HDI value experienced by that country. Hence, Figure 39 illustrates simultaneously, HDI values, HDi rankings, Changes in Values and Movement in Rankings.

Figure 40, on the other hand, shows changes in HDI rankings, together with HDI values and its components, for 35 African countries between 1970 and 2005. Here, the Development Tree makes it easy to compare changes in HDI and its components while arcs help connect countries that went up the ranking (green) or down the ranking (red). The use of the development tree in this figure creates a strong visual statement by providing a solid area comparison for the level of development of countries at different time periods. The substantially larger trees, representing the level of development of African countries in 2005 (right column), highlight an across the board increase in the level of Human Development for the entire region. A second feature that comes clearly out of Figure 40, albeit with more of a keen eye, is that in 2005 the largest contribution to HDI for most countries came from the education indicator. Indeed, it is relatively easy to notice that the largest contribution to HDI, shown by the rightmost box of the 2005 development trees, is given by the education indicator, which are represented by the blue boxes. Indeed, from the 35 countries in the sample, education was the largest contributor to HDI for 22 of them, whereas the Life indicator was the largest contributor to HDI for 12 of them with Botswana and Mali being the only countries in the sample for which GDP was the largest contributor to HDI.

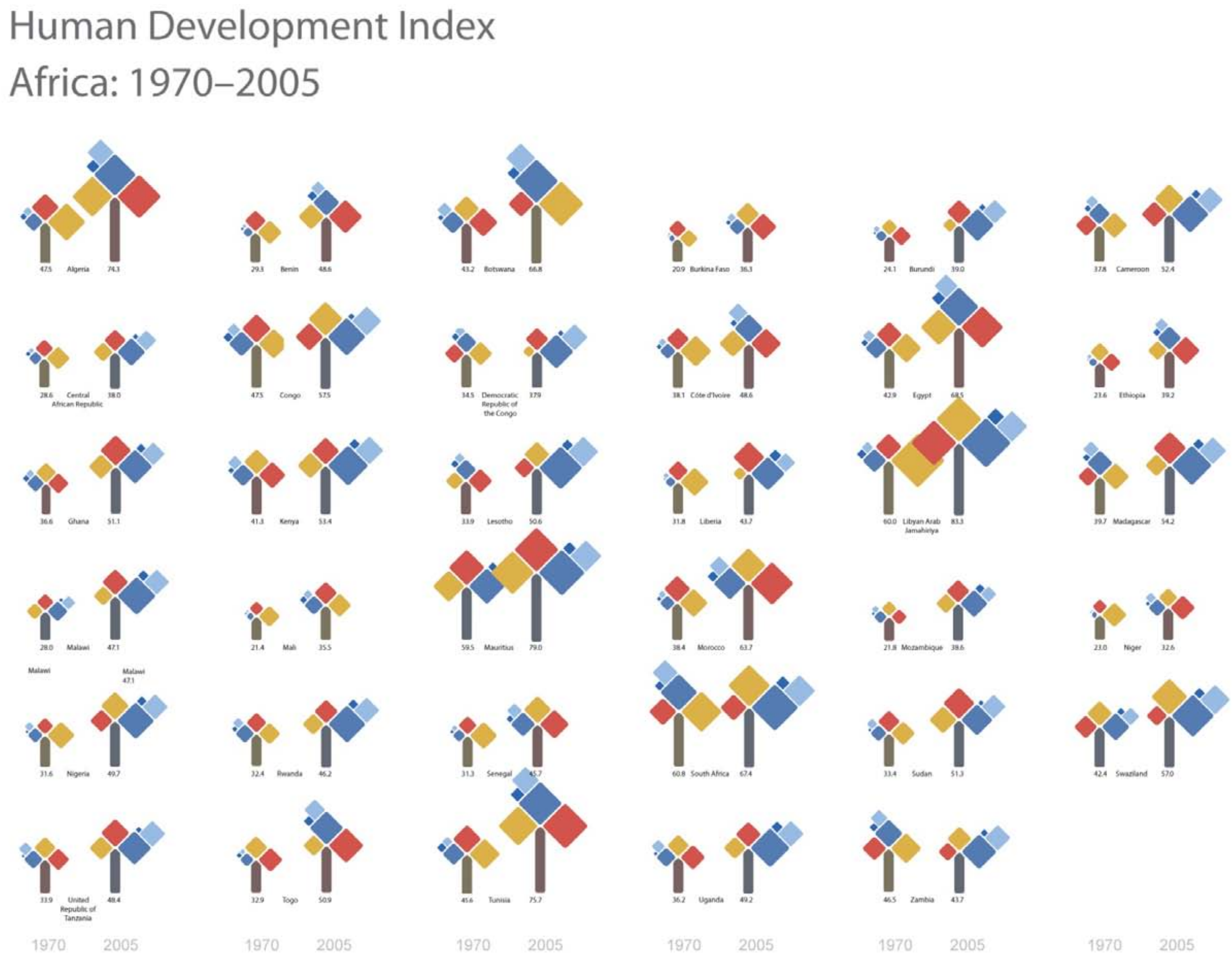

Figure 36 Comparison between the level of development, as measured by HDI and its components, using the Diamond Tree Design for 35 African Countries between 1970 and 2005.

Figure 37 Comparison between the level of development, as measured by HDI and its components, using the Spline Tree Design for 35 African Countries between 1970 and 2005.

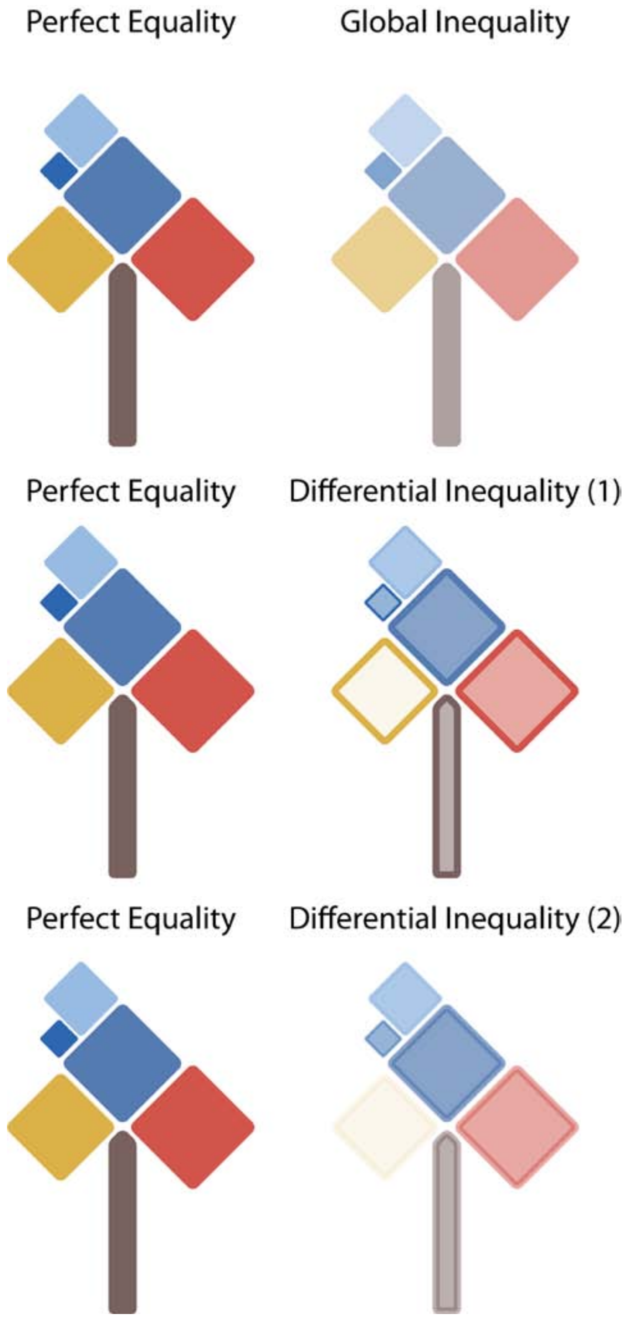

Figure 38 Example of ways in which the Diamond Tree Design can be used to visualize inequality together with the Human Development Index and its components. Here inequality is encoded through opacity with large levels of inequality being encoded by low levels of opacity. The first pair of trees (top) compares perfect equality (100% opacity) with a tree at 60% opacity. The middle and bottom designs show trees in which the opacity of each of its branches is different, encoding high levels of inequality for the GDP branch and mild levels of inequality for education (Figures are schematic, as no data on inequality was available at the time they were produced).

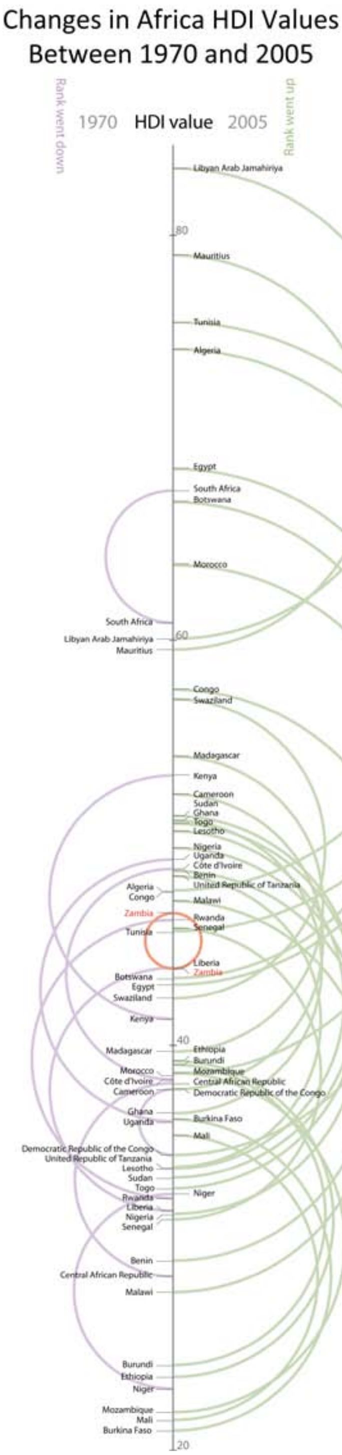

Figure 39 Comparison between HDI values in Africa between 1970 and 2005. All arcs, except, the red one (Zambia) show upward movements in HDI. The purple arcs in the left connect the countries who went down the ranking, despite their increases in HDI. The right arcs on the right connect countries who experienced an increase in both, HDI ranking and HDI values.

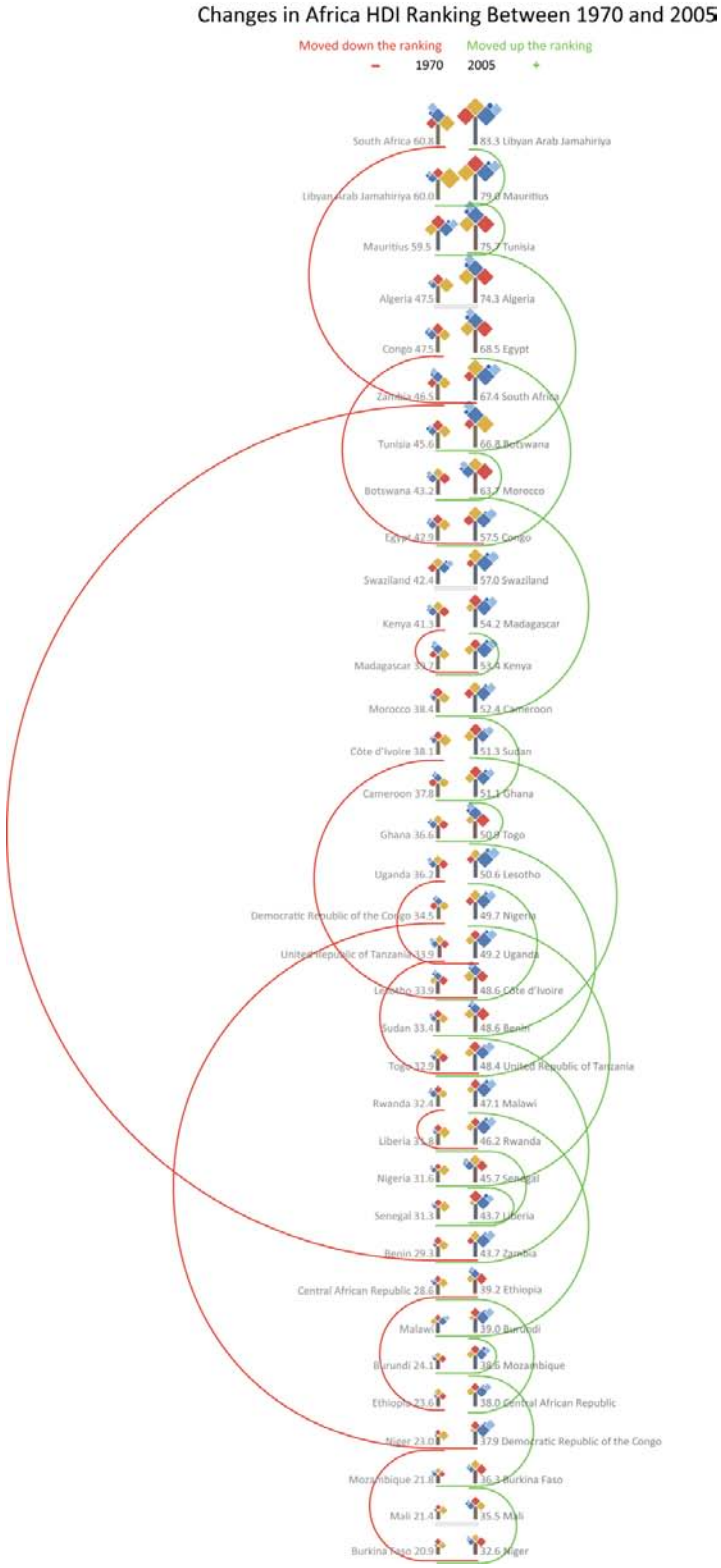

Figure 40 Comparison between the ranking of African countries in 1970 and 2005. Red arcs on the left connect countries that went down on the HDI ranking, whereas the green arcs on the right connect countries that moved up the HDI ranking. The values of HDI are represented by the tree and are also written numerically. The value of the components of HDI can be seen in the design of the trees.

(V) Partial Ordering Networks and the Development Reference Groups

While it is possible to sort and rank countries by their HDI values, it is important to understand that these ranking depends on the way in which the three components of HDI are aggregated. An alternative to the use of aggregation is to rank countries using partial ordering techniques that can be used to compare countries in absence of assumptions of the weights that should be assigned to each one of the three HDI components.

Here we introduce the idea of a Partial Ordering Network (PON) as a method that can be used to visualize and compare the relative level of development experienced by different countries. To understand how we define the PON consider the following examples, dealing with three countries, Argentina, Chile and Mexico.

COUNTRY	COUNTRY CODE	EDUCATIONx	LIFEx	GDPx
Argentina	ARG	0.95	0.83	0.78
Chile	CHL	0.92	0.89	0.80
Mexico	MEX	0.87	0.84	0.81

When we look at the value of the three indicators for these three different countries we find that the ordering that comes out of each one of them is different. If we consider the education indicator, for instance, we find that sorting countries from top to bottom results in the ordering ARG, CHL, MEX, whereas when we rank them according to GDP we obtain MEX, CHL, ARG. Moreover, ordering by the life indicator results in CHL, MEX, ARG. In this particular example there is no unique ordering that can be used between the level of development of these different countries and we could well say that ARG, CHL and MEX are at a comparable level of development.

If we would compare ARG, CHL and MEX, however, with Australia (EDx=1.04, LIFEx=0.93 and GDPx=0.96) we would find that no matter what dimension we choose, Australia will be located in the top of the ranking, and we could say that, no matter what dimension we consider more relevant, Australia is at a higher level of development than Argentina, Chile and Mexico. On the other hand, if we would repeat this exercise by comparing ARG, CHL and MEX with Ivory Coast (EDx=0.48, LIFEx=0.52 and GDPx=0.46) we would find that ARG, CHL and MEX score higher than Ivory Coast in all measures, suggesting that these three Latin American countries are at a higher level of development than Ivory Coast.

Since these relationships are not always transitive we define the Partial Ordering Network as a network connecting countries at a comparable level of development. Following the example presented above, Chile and Mexico are at a comparable level of development than Argentina, since Chile and Mexico score higher than Argentina in GDP, but lower than Argentina in education. In such cases we say that Chile and Mexico are in Argentina's Development Reference Group. Going forward we define a country's Development Reference Group (DRG) as the set of countries that score both, higher and lower in some of the HDI components than the country in question.

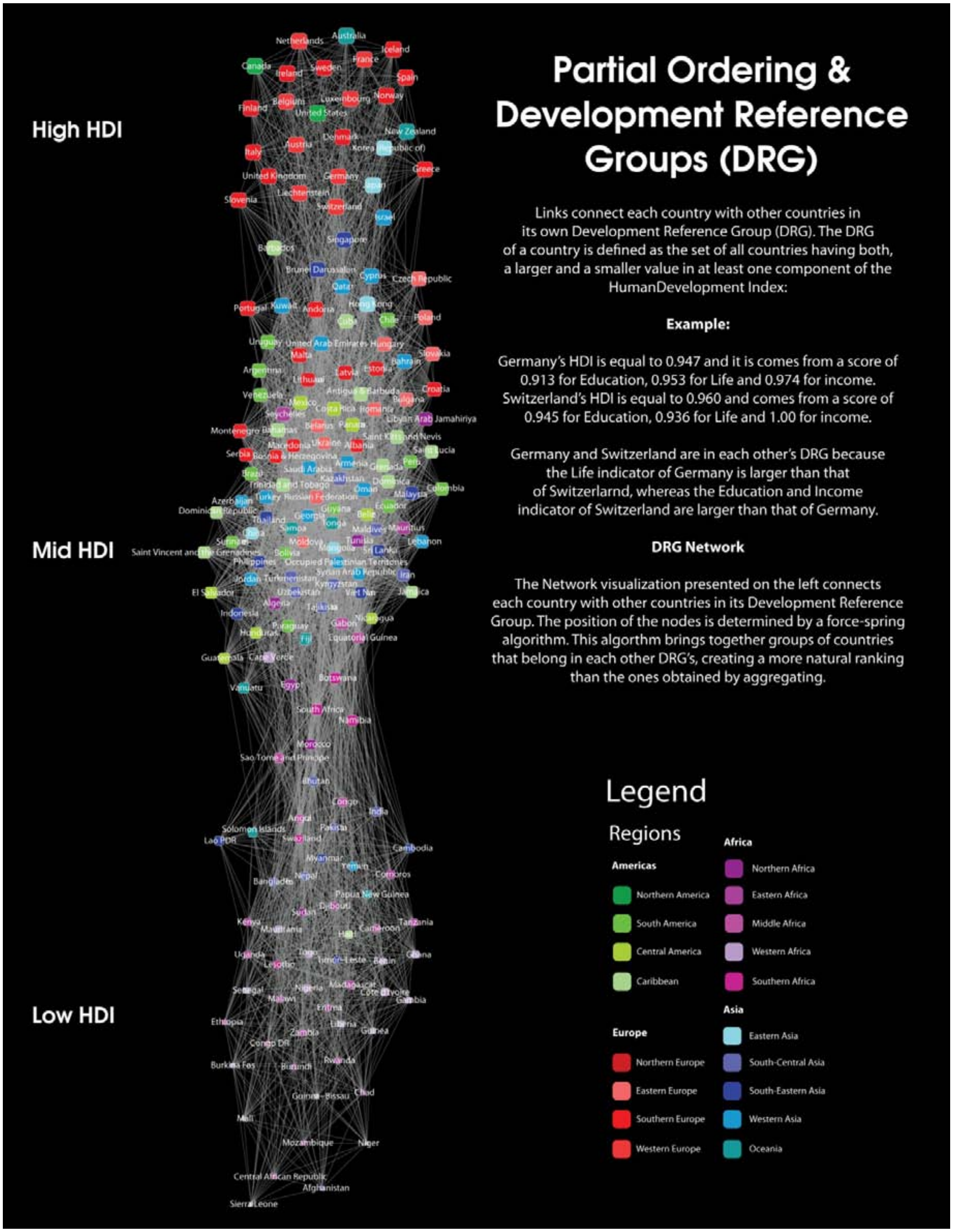

Figure 41 Partial Ordering Network created using the 2007-2009 estimates of HDI compiled by Gray and Purser.

Figure 41 shows a visualization of the Partial Ordering Network created using Gray and Purser's data for the 2007-2009 period. Each node represents a country and it is colored according the UN world macro regions and components that can be found at http://www.un.org/depts/dhl/maplib/worldregions.htm. Node sizes are proportional to HDI. Links connect countries in the same Development Reference Group, as explained above and in the figure. The network layout was determined using a force-spring algorithm which helps group countries with others in their same Development Reference Group.

The structure of the network shows the existence of three clearly distinct clusters of countries, a fact that is not clearly visible using other visualization techniques. Note that this natural groups are not the same as the Very High, High, Medium and Low groups that have been used in the report. At the top of the network we find countries from Europe, North America and a few from Asia and Oceania, such as Korea, Japan and Australia. The second cluster, which can be clearly appreciated further down the page, gathers most of the Americas, Asia, Eastern Europe and Northern Africa. Finally, after a considerable gap, we find most of Africa and parts of South Eastern Asia. The Partial Ordering Network provides a view of the level of development of different countries which is not constrained by the limitations of rankings and the clutter from single values, allowing us to better understand a country's level of development in the context of other countries in the world.

The Partial Ordering Network can be seen as a map that groups countries with its "development neighbors", rather than their geographic neighbors. For instance, Yemen, Papua New Guinea and Tanzania are countries at similar level of development and are therefore neighbors (connected by links) in the Partial Ordering Network.

It is also possible to illustrate the Partial Ordering Network that emerges from the three main HDI components by using a Geographic, rather than a relational layout (Figure 42). Figure 43 to Figure 46 show examples of such a visualization in which we have highlighted Chile and Haiti's Development Reference Group. This layout is very useful to understand the development context of a country and can be a useful technique to show how a nation compares to others, especially when these comparisons seem not to be very intuitive. For instance, in Chile's Development Reference Group we find Equatorial Guinea, a country that most Chileans (such as myself) would not identify, from the top of our heads, as a country with a high GDP per capita. Yet, the figure helps highlight this fact and helps, with additional references, inform of the relatively abundant oil exports coming out of the small East African nation. Examples like this one show how these visualization techniques can present relevant country specific information in ways which are accessible, visually attractive and informative for a wide audience. Making them ideal for museum exhibits and public campaigns.

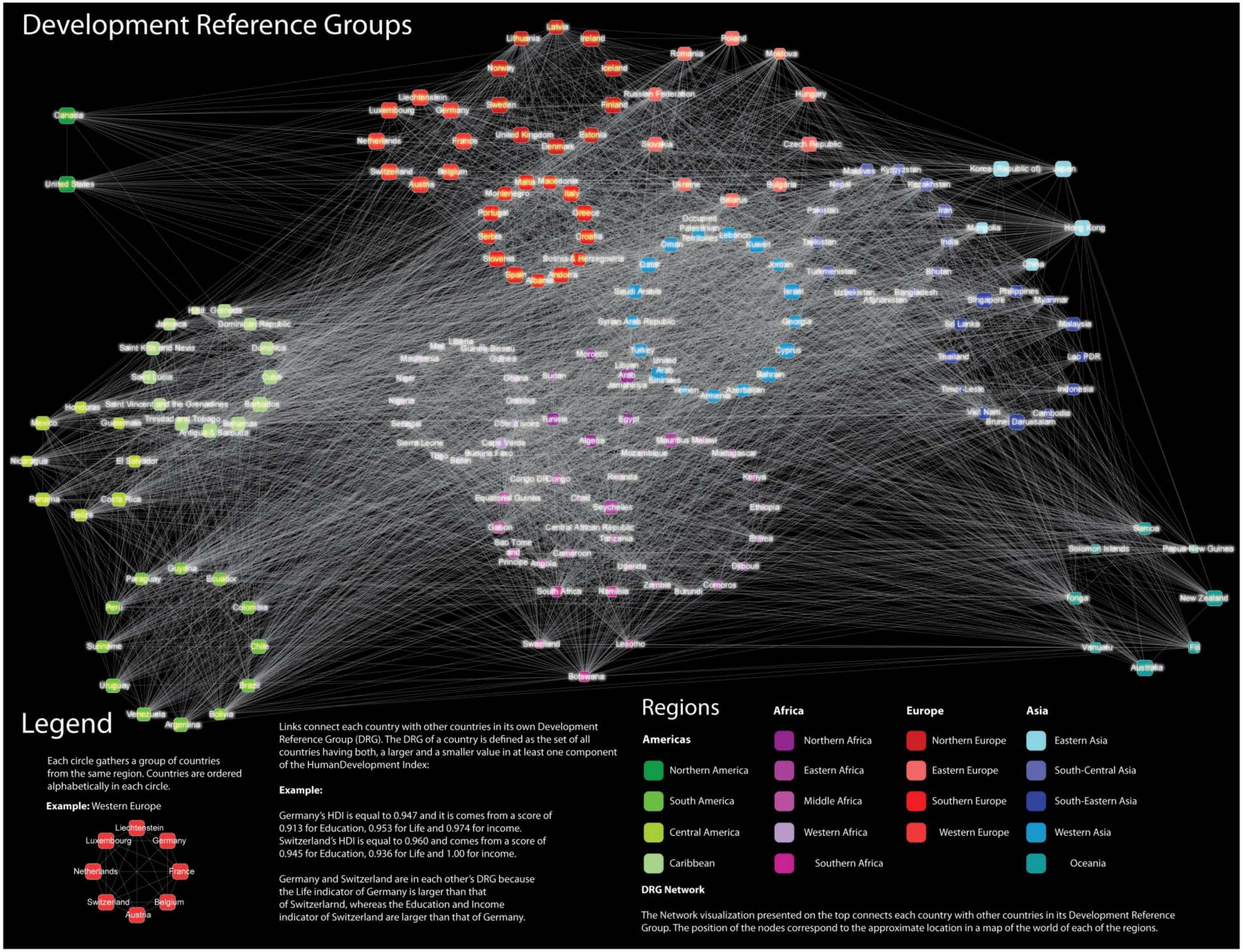

Figure 42 Partial Ordering Network using Geographical Layout

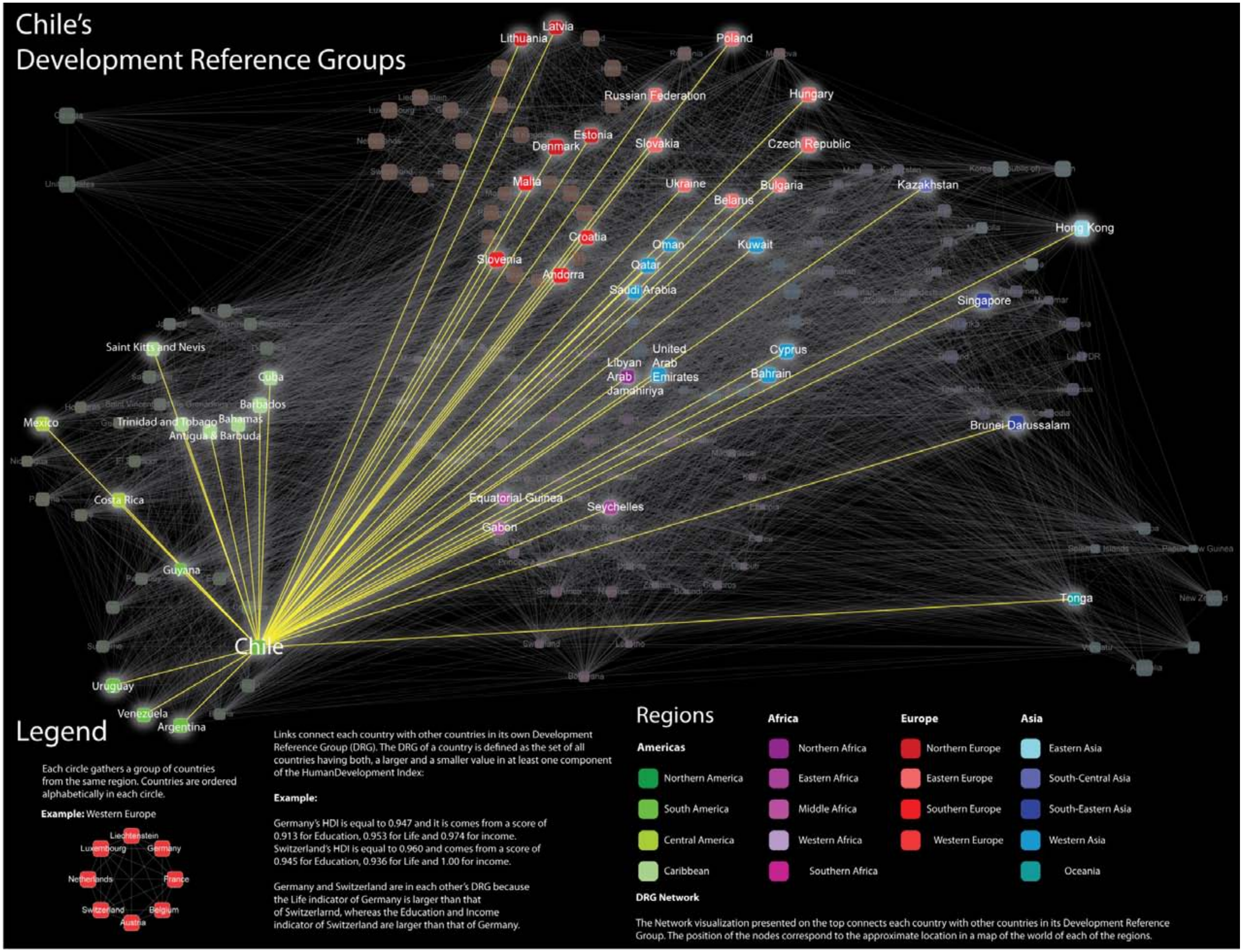

Figure 43 Chile's Development Reference Group. The position of a group of countries is determined by their geographical regions.

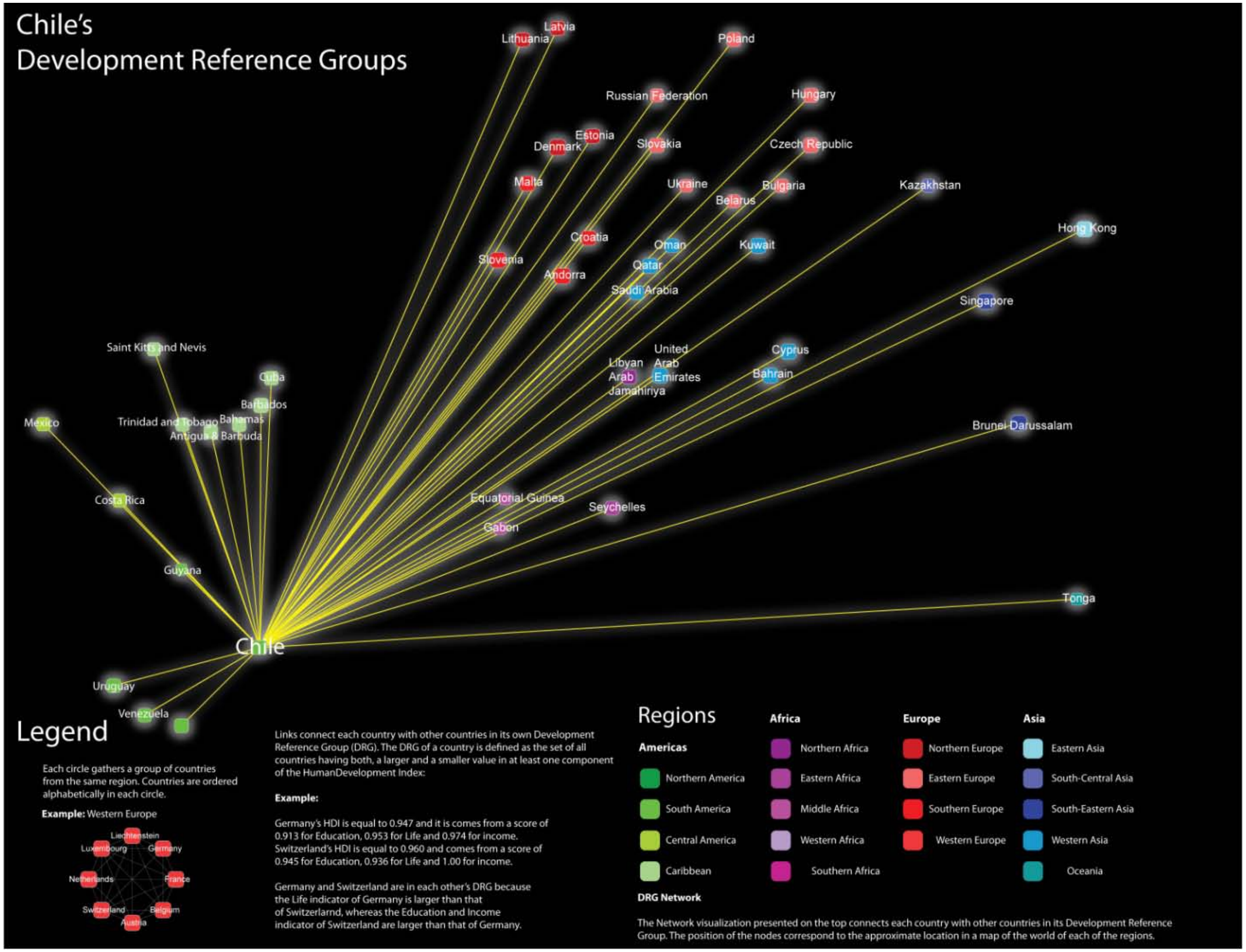

Figure 44 Chile's Development Reference Group using a Layout in which only countries in Chile's DRG are shown.

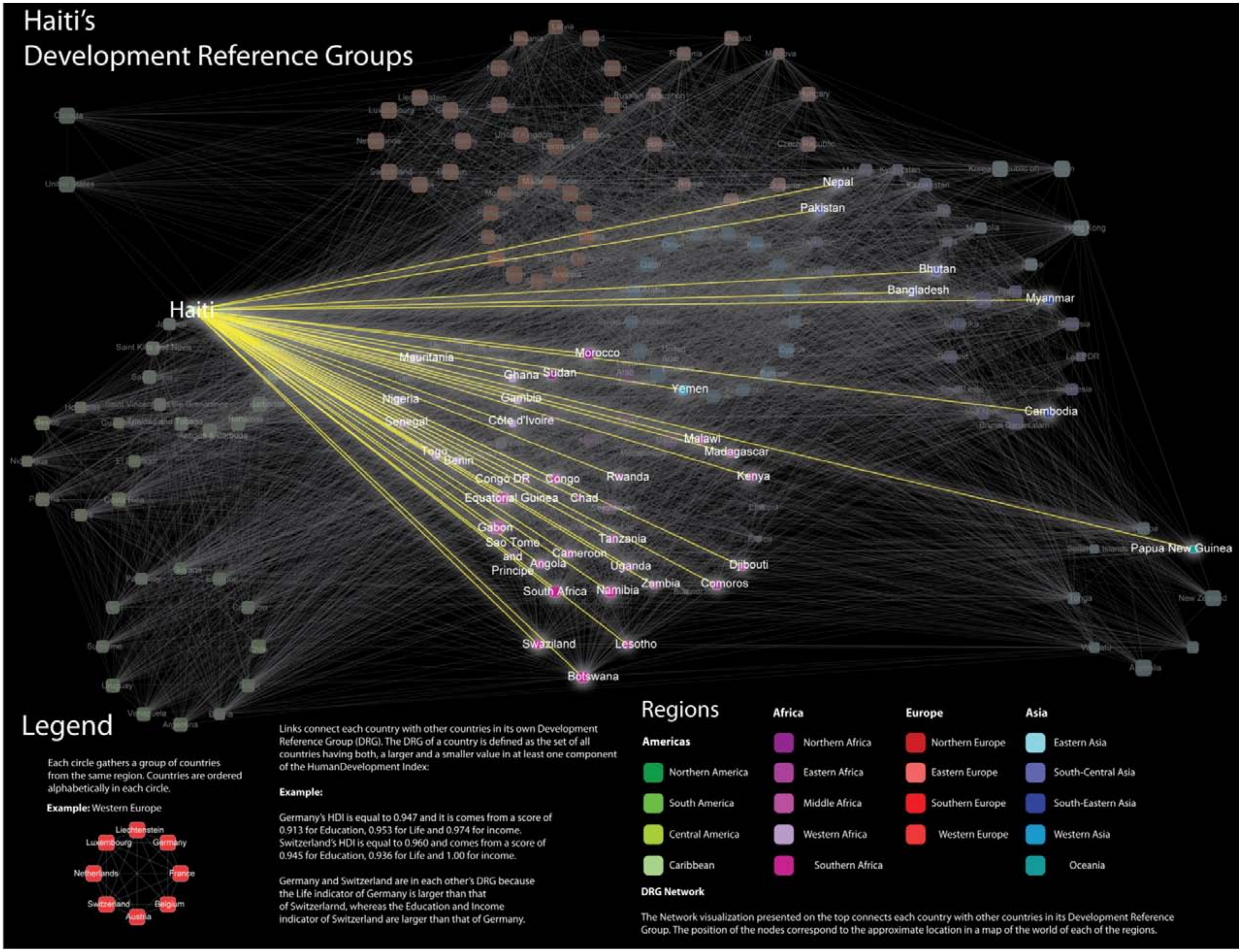

Figure 45 Haiti's Development Reference Group. The position of a group of countries is determined by their geographical regions.

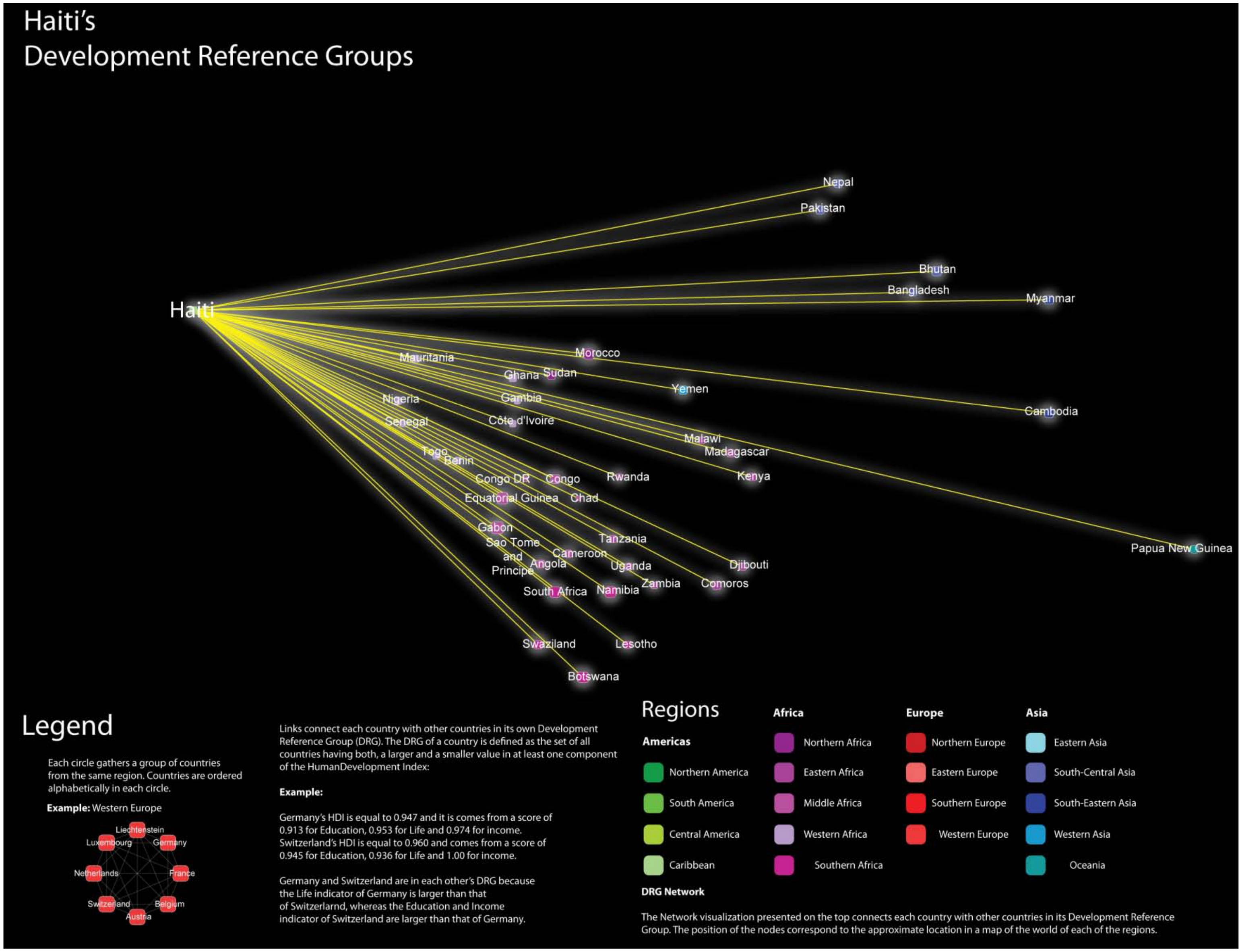

Figure 46 Haiti's Development Reference Group using a Layout in which only countries in Haiti's DRG are shown.

DISCUSSION

Here we have presented a visual exploration of HDI and its components across time through a myriad of representation techniques that allowed us to illustrate how HDI and its components have changed during a 35 year period. From a historical perspective, we saw that while the use of visualizations is not extremely prevalent in modern Social Sciences, it was developed in the past in the frontier between the Natural and the Social Sciences, by polymaths such as Joseph Priestly, William Playfair, Charles Joseph Minard, John Snow and Florence Nightingale.

Visualizations allow not only to represent and communicate information that can be expressed in narrative and numerical form, but also to highlight relationships that cannot be expressed using narrative and numerical techniques. An important aspect of visualizations is that they speak to a different part of our brains that written numerical and narrative descriptions, making them a more universal form of communication. Indeed, there is scientific evidence supporting the fact that graphical representations are more intuitive than more structured forms of language, like written and numerical language which require substantial learning. This is supported, for instance, by work done by Jan Deregowski during the 60's and 70's. Deregowski showed that Africans individuals with little or no exposure to graphical art were able to recognize objects from photographs and pictures (Dergowski 1972). This universal ability of interpreting graphical representations was also studied by Robert Serpell, who showed that urban American and rural Zambian children showed similar improvements in their progression on the ability to discriminate images, albeit with a 2-3 year lag in the developmental time (Serpell

1971). This suggests a universal physiological progression in the way in which we process graphical representations that appears to be independent of cultural constraints.

The existence of optical illusions also is evidence of the universality of graphical perception. Since optical illusions persist despite knowledge of the illusion itself, they show that there are purely biological aspects affecting our perception of graphical representations that cannot be changed by experience. Several examples of good optical illusions are nowadays widely available through the web. Here, for illustrations purposes only, we show an optical illusion created by Ewald Hering. Ewald Hering was an 18 century German physiologist that help developed the theory of color opponency, which is the basis used to explain and detect color blindness. Hering's illusion, shown in Figure 47, shows a couple of parallel red lines that do not seem parallel, but rather bowed outwards, when they are put in front of a particular array of lines that converge in the center of the image. The bottom panel shows the same two parallel lines in absence of the background lines that cause the illusion. The bottom panel of Figure 47 explains the illusion, yet it is not able to make it disappear.

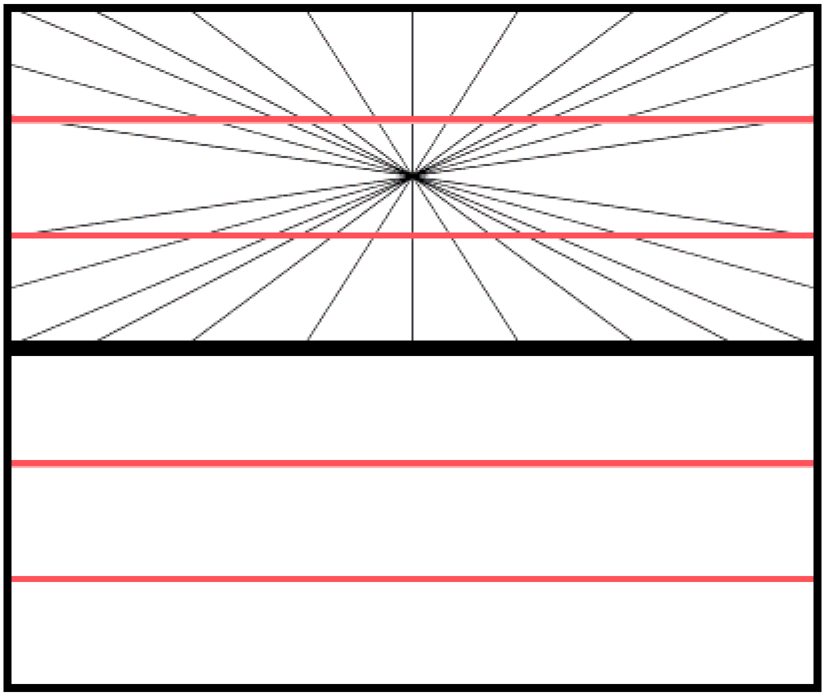

Figure 47 Herring illusion, after Ewald Hering (1861)

The visual exploration of Human Development performed in this study was limited by a relatively small dataset of Human Development and its components. Within these constraints several alternative for the visualization of the Human Development Index and its changes in time were explored. From these visualizations, the richest ones are those that include the Development Tree, which allows the simultaneous representation of the Human Development Index and its components, and the Partial Ordering Network, which allowed connecting countries to those at a comparable level of development. These designs can help in the creation of comparative charts which have not only a numerical, but also a symbolic and emotional content. We strongly suggest the use of the Development Tree as a potential graphical branding of the Human Development Index and for use in educational and outreach campaigns.

ACKNOWLEDGEMENTS

I would like to acknowledge the enormous help provided by Isabel Meirellis, Alex Simoes, Dave Landry and Geoff House in the design of the Development Tree. They deserve most of the credit for this. I would also like to acknowledge the support of Francisco Rodriguez, Jeni Klugman, Jean-Yves Hamel, Isabel Pereira, José Pineda and all the members of the Human Development Report office at UNDP for useful discussions and comments throughout the creation of this manuscript.

REFERENCES:

L. Brasseur, Florence Nightingale's Visual Rhetoric in the Rose Diagrams. *Technical Communication Quarterly* 14(2) 161-182 (2005)

JB Dergowski, E S Muldrow, W F Muldrom, Pictorial recognition in a remote Ethiopian population, Perception 1(4) 417-425 (1972)

Gaulier et Zignago "BACI: International Trade Database at the Product-level The 1994-2007 Version" Cepii Working Paper (2009)

S. Hempel, The Strange Case of the Broad Street Pump: John Snow and the Mystery of Cholera. University of California Press, Berkeley, CA (2007)

C.A. Hidalgo, B. Klinger, A.-L. Barabási, R. Hausmann, The Product Space Conditions the Development of Nation, *Science* 317:482-487 (2007)

George Gray Molina, Mark Purser, Human Development Trends since 1970: A Macro Story of Micro Changes, HDRO Working Paper Paper 01/2010, New York: UNDP (2009)

R Serpell, Discrimination of orientation in Zambian children. Journal of Comparative and Physiological Psychology 75(2) 312-316 (1971)

ER Tufte, Beautiful Evidence, Graphics Press, Cheshire, CT (2006)