

Case Study: The Development of the HDI Tree

Isabel Meirelles¹, César A. Hidalgo^{2,3}, Alex Simoes²

Design process, graphical systems, metaphors

This paper describes the process of developing graphical metaphors and diagrammatic representations for the Human Development Index – HDI— and its components: the HDI Tree (UNDP, 2011). In 1990, the United Nations Development Programme introduced the Human Development Index as a measure of development that combines indicators of life expectancy, educational attainment and income. The project was commissioned to César A. Hidalgo (2010) by the Human Development Report office and developed in collaboration with three senior students at Northeastern University. The objective was to explore ways to simplify and communicate the Human Development Index using visual rather than numerical representations. Here, we describe the design process, discuss the concepts behind the metaphor, and analyze two selected visualizations. We conclude by examining how the Development Tree can be used in visual narratives for educational and outreach purposes

Introduction

In 2009, the Human Development Report Office invited César A. Hidalgo to write a research paper examining graphical alternatives for the Human Development Index. In the background paper, Hidalgo introduced “five graphical statistical methods to compare countries level of development relative to other countries and across time. For this, seven panels of data on the Human Development Index and its components were used. These contained information on more than 100 countries for more than 35 years. The five graphical statistical methods represented: (i) Rankings (ii) Values (iii) Distributions (iv) visual metaphors (The Development Tree), and (v) Partial Ordering Networks (PON) which were used to introduced the concept of Development Reference Groups (DRG)” (Hidalgo, 2010, abstract).

This paper describes the design process by which the HDI tree was developed. The work was performed together with Northeastern University students during Spring 2010, prior to their graduation: Geoff House, David Landry and Alex Simoes.

What is the HDI?

In 1990, the United Nations Development Programme (UNDP) launched the *Human Development Reports* (HDR) and introduced the Human Development Index (HDI). Pakistani economist Mahbub ul Haq together with Nobel laureate Amartya Sen and other leading thinkers developed the HDI “as an alternative to conventional measures of national development, such as level of income and the rate of economic growth.” (HDR, retrieved 07.28.11). By complementing economic values with achievements in health and education the HDI aimed to emphasize the human capabilities present in a country with the goal of steering policy to objectives related to human development.

The HDI is a composite index that aggregates three dimensions: health—measured by life expectancy at birth; education—measured by a combination of both, adult literacy and primary school enrolment; and income—measured by purchasing-power-adjusted per capita Gross Domestic Product (GDP) (PPP US\$). In 2010, for the 20th anniversary edition, the HDI replaced

¹ Department of Art and Design, Northeastern University

² Macro Connections, The MIT Media Lab, Massachusetts Institute of Technology (MIT)

³ Center for International Development, Harvard University

the education indicator with (i) Expected years of schooling and (ii) mean years of schooling. The formula used to calculate HDI, was also changed from arithmetic to a geometric mean. Here, however, we look at visualizations constructed with the initial definition of the HDI, as this was the official version at the time the work was commissioned, the only exception being the application of the methods to the creation of an Online Tool.

The motivation to introduce graphical representations is that two countries with similar HDI values might differ in the value of the HDI components. As an example, we compare Brazil and Georgia, who have close values in HDI despite having large differences in income and education.

BRAZIL		GEORGIA	Notes
67.80	HDI	67.90	Almost the same HDI values
61.71	Income	48.71	Income values are quite different
81.80	Health	81.86	Health indicators are almost equal
61.65	Education	78.68	Large differences in education
\$8,982.00	GNI per capita [PPP USD]	\$3,856.00	Georgia GNI is less than half
72.00	Life Expectancy [Years]	72.00	Exact same life expectancy values
7.00	Mean Years of Schooling	7.00	
14.00	Expected Years of Schooling	13.00	

As with any index, HDI is a crude measure the development capabilities of a country. The goal of this paper, however, is not to discuss the merit or value of the index, but rather to present the process of visually representing it.

The HDI numerically

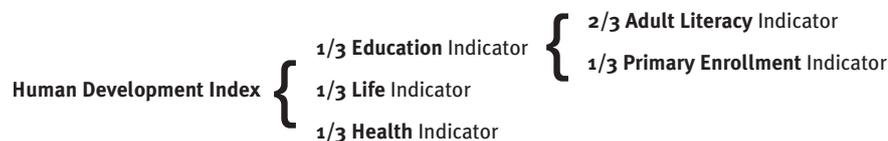
The Human Development Index (HDI) is a statistic used to describe countries well being that gives equal weight to (i) an education indicator, (ii) a life or health indicator, and (iii) an income indicator. Formally the HDI of country c at time t is defined as:

$$HDI_c(t) = \frac{1}{3}ED_c(t) + \frac{1}{3}LIFE_c(t) + \frac{1}{3}GDP_c(t)$$

Where $ED_c(t)$, $LIFE_c(t)$ and $GDP_c(t)$ represent, respectively, an education indicator, which combines (2/3) adult literacy and (1/3) primary school enrolment, 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) adjusted by purchasing power parity. All of these indicators are normalized such that their values lie between 0 and 1.

The HDI schematically

The HDI can be schematically represented as the hierarchical structure below:



We often describe this type of schema as a *hierarchical tree*. Thus, we could as well read the schema above as a tree structure where the HDI value is the trunk and its indicators (components) are the branches. Remind that our task was to design a visual display for representing the HDI as an alternative to the mathematical forms currently used to communicate the index. The main objective was that it would serve as an educational tool to provide practical guidance and help promote development among the general public across different countries in the world. Similarly to dietary visual displays around the world, such as the US food pyramid, our goal was to devise a diagrammatic representation that, at a glance, could communicate the HDI and its components (for a discussion of the use of metaphors in health displays see Meirelles, 2007).

The question that we explore here is the use of the *tree* as a *conceptual metaphor* to synthesize and visually represent the HDI and its components. Can we eschew from numerical representations and effectively communicate the HDI data to a general audience using the *tree* metaphor?

Metaphors

Metaphor is used in this paper as a “cross-domain mapping in the conceptual system” (Lakoff & Johnson, 1980; Lakoff, 1987). In *Metaphors We Live By* Lakoff and Johnson (1980) contend that abstract concepts are metaphorically understood in terms of more concrete and typically spatial concepts. They describe the role of metaphors as a conceptual mechanism in which in order to grasp concepts that are abstract or not clearly delineated in experience we use other concepts that are clear to us such as spatial orientations, physical experiences, known objects, etc. Metaphors are part of our conceptualizing processes and help us understand one domain of experience in terms of another.

The HDI data are not inherently visible, since the index is a statistic provided by the aggregation of three indicators, which are themselves abstract. For instance, life expectancy is not a tangible item, but an abstraction providing us with the potential longevity for a given population under a set of conditions. The same could be said of the education indicator. The aggregated HDI itself, however, is a more abstract concept that mixes these seemingly distinct indicators through a predefined mathematical formula.

Lakoff suggests that a familiar image-schema with a well-understood structure, can be used to metaphorically organize other complex concepts. Furthermore, his Spatialization of Form hypothesis maintains that “hierarchical structure is understood in terms of PART-WHOLE schemas and UP-DOWN schemas (Lakoff, 1987, p. 283). An instance of the part-whole schema is the way in which we experience our bodies—as wholes with parts—and how we use the body to explain other concepts. Consider the metaphor “company is body” and descriptions such as, the president is the head, the board is the heart, and the workers are the arms and legs, and so on.

The use of the tree metaphor to visualize data is not new. Trees or *arborea* are commonly found in medieval manuscripts illustrating all sorts of disciplines: “...these *arborea* had a solid authoritative base in the Bible. *The Book of Genesis* speaks (II, 9, 17) of the Tree of Life (*lignum vitae*) planted by God in the Garden of Eden and the Tree of Knowledge of Good and Evil (*lignum scientiae boni et mali*), the eating of whose fruit led to the expulsion of the Garden” (Murdoch, 1984, p. 38). According to Murdoch one of the earliest medieval uses of the *arbor* in a written text was to represent genealogy and can be found in manuscripts of the *Liber etymologiarum sive originum* of the seventh-century bishop Isidore of Seville (1984, p.47). Examples of *arborea* to structure other subject abound: from medieval dichotomies to Darwin’s original sketch of his evolution theory (not illustrated here due to copyright issues).

Even though we now use the tree schema in most representations of genealogy, such that we commonly call them “genealogical trees,” it is interesting to note that the human body preceded as the image-schema for structuring such data in medieval times as Klapisch-Zuber writes:

With regard to the representation of genealogical ties, the house or the human body often proposed their structure, and here and there the “branches” evoked the tree, the leaves and flowers. Between the ninth and twelfth century, all these schemas were tested without the selection of one in particular. They were even combined so as to accumulate the greatest number of possible directions. The clarification and focus on the image of the plant started from the twelfth to thirteenth centuries, at the end of a long meditation of the monks and Masters of the schools, on the symbolic values attached to the growth and flowering or the fruiting. Here, the tree was unbeatable. Springing towards the sky, most of its foliage and flowers reappearing each spring, safely carrying fruits, it nourished the imagination, its metaphor constantly enriched, widely commented on, supplied the representation of kinship, of consanguinity, of genealogy. If it were a question of visually restoring the idea of fertility, progeny, existential strength through the ages, then the plant had no rivals. (2003, p. 37; my translation, original text in References)

The HDI metaphorically

For the HDI project, we considered the metaphor: *development is growth*, in which we use the *tree* as the source part-whole configuration for structuring the target concept of the HDI and its

components. Trees are well-known living organisms familiar to all peoples of all countries independent of their culture or flora. We all understand trees as formed by parts: roots, trunk, branches, leaves, etc. Similarly, we all understand the development aspect of such omnipresent type of plant. It could also be argued that the ubiquitous use of *arborea* to represent all sorts of subject, for over twelve thousand years, and across the world, the *tree schema* can be considered a *familiar schema* to a large number of people.

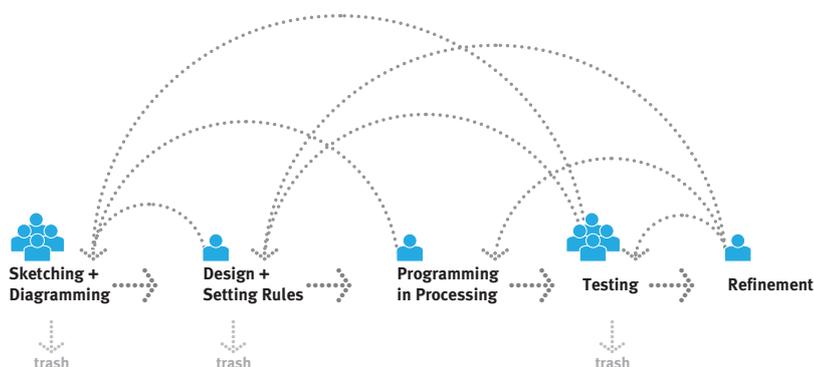
Research on the cognitive operations executed by a person in the process of reading quantitative graphs indicates that “people create schemas for specific types of graphs using a *general graph schema*, embodying their knowledge of what graphs are for and how they are interpreted in general” (Pinker, 1990, p. 104, italics in original). The pervasiveness of trees as a structuring schema is fundamental to our project, in that, our goal was to provide comprehension of a less familiar index to the general public. The familiarity with the tree schema might help the reader in mediating between perception and memory, and ultimately, might facilitate understanding.

In a diagrammatic representation the graphic elements and the graphic structure stand for elements and relations in another domain. Meaning is conveyed by means of visual references. Efficiency in conveying meaning will depend on how the visual description stands for the content being depicted, whether the correspondences are well defined, reliable, readily recognizable, and easy to learn (Pinker, 1990).

Finally, the choice of a part-whole schema allows the visualization of how individual indicators contribute to the whole index, thus providing a richer image of the HDI in contrast to a single numerical indicator. The structural elements of a part-whole schema are: a whole, parts, and a configuration. The configuration is a crucial structuring factor in the part-whole schema. Because the parts can exist without constituting a whole, it is the configuration that makes it an image-schema. Lakoff explains that “we have general capacities for dealing with part-whole structure in real world objects via gestalt perception, motor movement, and the formation of rich mental images. These impose a pre-conceptual structure on our experience” (1987, p. 270).

The design process

With a clear understanding of the goals of the project, the selection of a metaphor and an image-schema to structure the data, as well as an understanding of the main audience, we set up to explore design possibilities. The work was developed over 45 days with a team of four people. Considering the short time for developing the project—from concept to implementation—we devised a process in which we worked both as a group, mainly for brainstorming, critiques and testing, and individually or in small groups to perform specific tasks and the bulk of the work. Due to an effective group dynamic we were able to arrive at solutions that we feel were authored by the whole and not by the parts. The diagram below describes the main activities and the people involved.



Simultaneously to initial sketching, we set up a database of the HDI data, which comprised information on more than 100 countries for 35 years. The project uses the HDI data compiled by Gray and Purser (2009) summarizing HDI and its Components between 1970 and 2005 in five-year intervals, where data availability determines HDI country coverage. Programming was developed using the open source Processing environment.

Spline Design Rules (fig. 2)

- The height of the trunk is linearly proportional to the HDI value
- 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 to 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.
- The point at which the branches begin (change of colour) is proportional to its contribution to the total HDI.
- The length of the branch (after the curve) is proportional to the actual value of the component indicator.
- The colour of the trunk is a weighted average of the colour of the HDI components.

The Pros:

- Branches are encoded by linear shapes thus reinforcing the linear dimension of the indicators
- The spatial position of branches according to ascending order helps inference of which components contributed the most to the HDI value.
- Having the colour of the trunk as a result of the weighted average of the colours of the components provides a double code, which reinforces information of how much each component contributed to the whole HDI. (Fig. 3)

The Cons:

- Non-contiguous location of branches makes comparison between indicators harder to perceive, and inference of the relative differences in values is onerous.
- The spatial position of branches according to ascending order impairs detection of indicators when making comparisons due to placement changes (see further discussion related to OPTION 2 below)

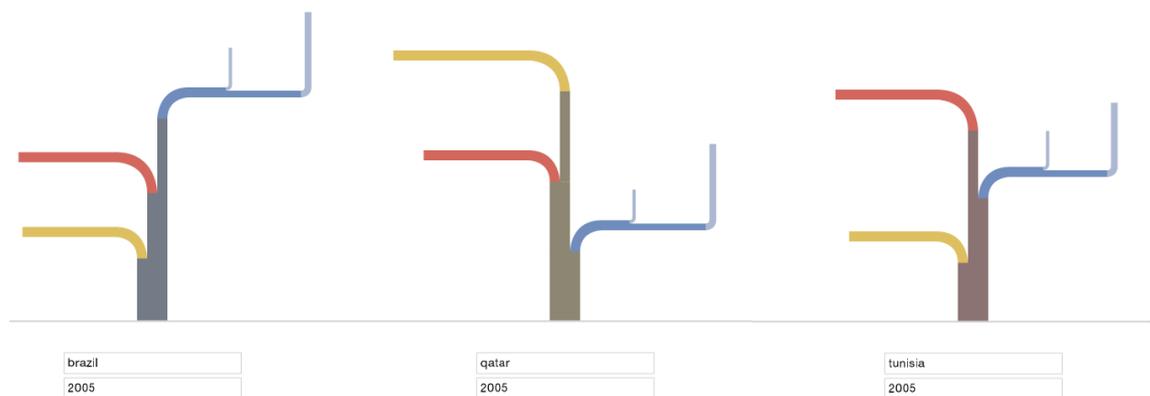


Figure 3: Illustration of how colour of the trunk reflects the composition of indicators (used with permission of Hidalgo).

Option 2: Diamond Development Tree

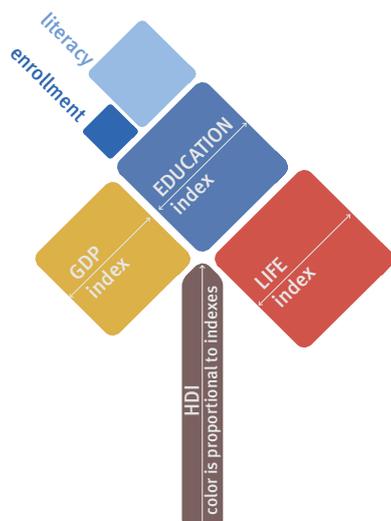


Figure 4: Diamond Design Rules (used with permission of Hidalgo)

Diamond Design Rules (Fig. 4)

- The height of the trunk is linearly proportional to the HDI value
- The order in which the branches come out from the tree indicates the relative contribution of each component to HDI. The left most branch being the smallest relative contribution to the value of HDI and the right most diamond representing the highest relative contribution.
- The side of the diamond is proportional to the actual value of the component indicator.
- The colour of the trunk is a weighted average of the colour of the HDI components.

The Pros:

- Contiguous location of branches facilitates comparison between indicators, making the relative differences in values readily available and easy to perceive.
- The spatial position of branches according to ascending order helps inference of which components contributed the most to the HDI value. (Fig. 5)
- Having the colour of the trunk as a result of the weighted average of the colours of the components provides a double code, which reinforces information of how much each component contributed to the whole HDI.

The Cons:

- The spatial position of branches according to ascending order impairs detection of indicators when making comparisons due to placement changes. This item was criticized at the UNDP web log and discussed below (Fig. 5).
- Branches are encoded by diamond shapes, which might be confusing considering that they encode linear dimensions.

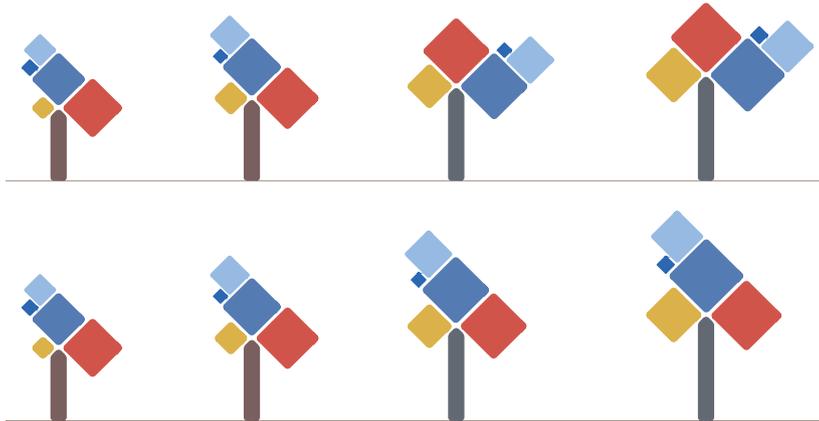


Figure 5: Series of Development Trees of China for years 1975, 1985, 1995 and 2005. The top row shows the ascending order and in the one below placement of the indicators is kept constant (used with permission of Hidalgo).

Online Tool

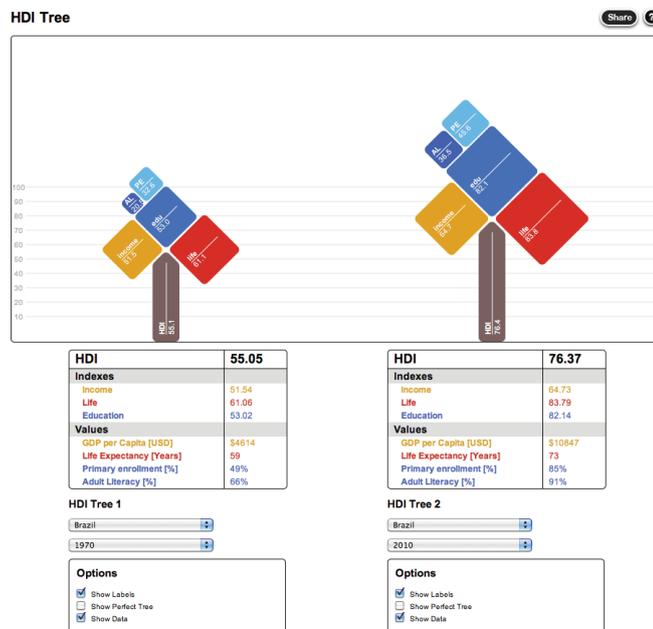


Figure 6: Screen shot of online HDI applet <http://hdr.undp.org/external/hdi/index.php> (used with permission of Hidalgo)

In November 2010, Alex Simoes implemented the Diamond Development Tree as a JavaScript online tool (Fig. 6). It first appeared at *Forbes Online* (2010) and later at the Human Development Report office website from UNDP (2011). The application uses the hybrid HDI data introduced in the 2010 report provided by the HDR office at the UNDP (REF). More specifically, the application considers the dimensions used to calculate HDI values up to 2009 as the indicators and the formula introduced in the 2010 report to aggregate these values into the HDI. The period covered is between 1970 and 2010.

Viewers Criticism

A total of 36 comments were posted at the UNDP site (data last collected on June 30th). Twenty percent of the comments criticize the increasing criteria for the ordering of indicators and suggest that a consistent positioning would have made comparisons across countries and/or years easier. We cite below two commentators.

David Hastings, Curator, Human Security Index (humansecurityindex.org) :

“The idea of a graphical representation can be useful. But may I suggest that the proposed tree be slightly modified to facilitate skimming, as well as more intuitive assessment? If you keep each elemental branch (education, health, income) in the same position for all countries, but merely adjust the size of each branch as you do, the tree is more intuitively analytical. The current re-positioning according to size impedes assessment (maybe not for the original author, but for many others).”

Luis Eduardo Guarnizo, Professor University of California, Davis:

“I think this is a very significant contribution. However, I'd suggest not to order the branches in increasing order, for it makes the visual comparison awkward. If you keep the tree design fixed, then it would be much easier to compare countries by size & colour simultaneously. When the location of the boxes changes, it becomes visually dissonant, adding an unnecessary step in the comparison.”

When writing the PROS and CONS of the design solutions we mentioned the position of branches in both sides of the critique. It is our view that each way of ordering actually answers a different question and highlights different aspects of the index:

1. **Ascending ordering** answers the question of which item weighted more in composing the index. Because the result stresses the relative amount of contribution by each indicator in the overall composition, it is easier to infer what indicator contributed the most or the least. The understanding of relative values is facilitated whereas comparison between the same indicator across countries or years is onerous. For example, countries with the same HDI might have different composition of the categories (Fig. 7). Also in an analysis of a particular country over time, it is possible to examine how the categories changed (if at all) over time (Fig. 8). Because numbers are very similar in amount, the ordering helps to quickly grasp the indicators with larger impact in the overall index.
2. **Consistent ordering** answers the question of what is the overall composition of the index. Because the positioning of indicators is constant, it is easier to compare the same indicator across countries or years without having to look for a new position as in the re-ordering version. It makes comparison easier but inference of relative values is impaired (Fig.9).



Figure 7: Comparison of four different countries with similar values for HDI in 2010. Note how indicators impact the overall index (used with permission of Hidalgo).



Figure 8: The image illustrates Zambia from 1980-2010 in 5-year intervals. Note how indicators have changed over time including their weight in the overall HDI composition. (used with permission of Hidalgo)



Figure 9: This image shows same data as Figure 8 —Zambia from 1980–2010— but with constant placement for the indicators as a comparison to forge further discussion —on the placement criteria. (used with permission of Hidalgo)

Worth noting that, Maywa Montenegro’s article at the *Visualizing.org* web-site explicitly compliments the fact that the order changes: “Notice how the designers have sorted the three HDI components left to right, in ascending order, so it’s easy for us to see what factors have the most influence.” (Montenegro, 2011)

Independent of what side of the discussion one stands, it is relevant to remind that it is through discrimination (same-different dichotomy) in early stages of object perception that elements and patterns are detected and ordered. Patterns are central to how visual information is structured and organized. Literature in perception and graphical methods (Cleveland, 1994; Kosslyn, 1994, 2006; Ware, 2004) explains that it is easier to detect patterns if categories are ordered. This principle complies to Lakoff’s Form Hypothesis that linear quantity scales are understood in terms of linear order schemas (Lakoff, 1987).

All in all, the great majority of comments posted at the UNDP were complimentary, even when making suggestions such as those related to the ordering criteria as described above. Below we provide two examples:

Neda Jafar, Statistician, Gender and MDGs United Nations ESCWA:

“Congratulations, this is really innovative and well structured easy to grasp. I would like to have all Arab countries trees if possible.”

Naomi Black, Professor Emerita, Political Science and Women’s Studies York University:

“I think this is a splendid idea. I’m particularly interested in its ability to emphasize vividly the components of the indexes. People tend to get transfixed by the numbers and the rankings. ...”

Discussion

The paper described the design process of devising visual representations as an alternative to the numerical representation currently used to communicate the HDI and its components. The HDI is a composite measure of one health, one income and two education indicators, which are aggregated numerically through a set of formulas. The goal was to provide an educational tool to inform the general public about human development capabilities. For that we used the metaphor of *development is growth*, and the image-schema of the *tree* as the part-whole configuration for structuring the HDI data in a way that would be easier to understand by a wide and culturally diverse audience. We described two design solutions including a brief critique of their perceptual and cognitive potentials and failures. We understand that none of the proposed Development Trees are entirely effective to provide all possible information on the HDI. However, the visualizations offer the potential of conveying complex abstract information in a synthetic manner. It is true that by doing so, details are lost, since they become snap-shots of a larger issue, thus only conveying main concepts. On the other hand, the proposed Development Trees allow for meaningful graphical narratives such as the one presented in the figure below, where it is possible to quickly grasp information ranging from the contribution of components in a single country, the development of a country over time, to comparisons of countries in a continent (Fig. 10). Can it be used to affect policy making? Most probably not, but it certainly offers the potential for outreach and human development awareness campaign to non-technical audiences.

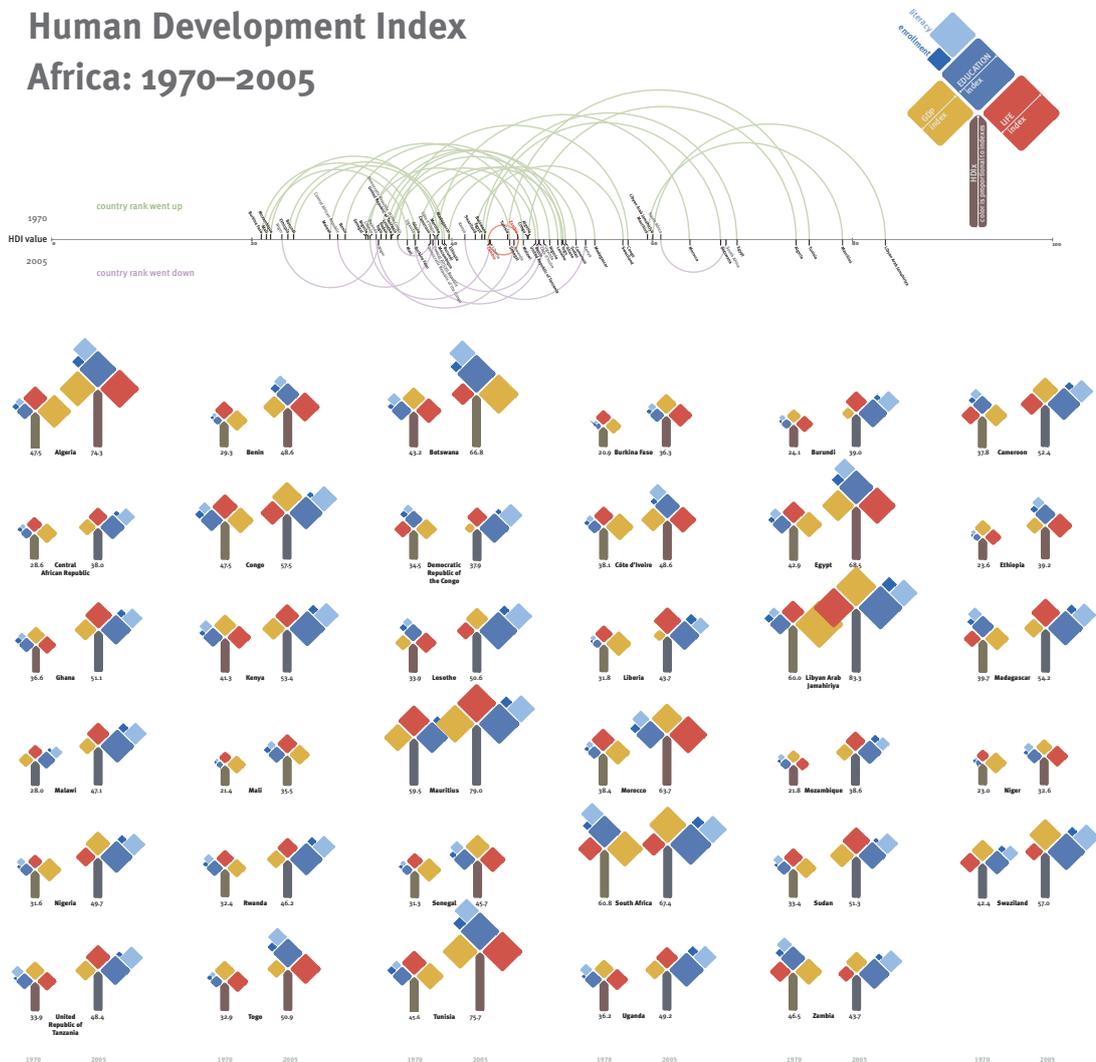


Figure 10: Poster designed to demonstrate the power of a graphical narrative with the purpose of discussing the HDI in African countries over time (used with permission of Hidalgo).

Acknowledgment

We acknowledge the comments of and support of Jean-Yves Hamel, Jeni Klugman and Francisco Rodriguez from UNDP's HDRO.

References

- Cleveland, W. S. (1994). *The Elements of Graphing Data*. Murray Hill, NJ: AT&T Bell Lab.
- Gray Molina, George & Mark Purser. (2009). "Human Development Trends since 1970: A Macro Story of Micro Changes", *HDRO Working Paper 01/2010*, New York: UNDP.
- Hidalgo, C. A. (2010). "Graphical Statistical Methods for the Representation of the Human Development Index and its Components," *Human Development Research Paper 2010/39* <http://hdr.undp.org/en/reports/global/hdr2010/papers/HDRP_2010_39.pdf>
- Hidalgo, C. A. (2010). "The Visual Human Development Index: Introduction", *Forbes Online* <<http://www.forbes.com/2010/11/17/human-development-index-visualization-hidalgo.html>>
- Hidalgo, C.A., Meirelles, I. & Simoes, A. (2011). "The HDI Tree: A Visual Representation", in *Let's Talk Human Development* (blog), web-site of HDR-UNDP. <<http://hdr.undp.org/en/humandev/lets-talk-hd/2011-02> >
- Human Development Reports-HDR (retrieved 07.28.11) "The Human Development Index (HDI)" in *HDR-UNDP* web-site <<http://hdr.undp.org/en/statistics/hdi/>>
- Klapisch-Zuber, Christiane. (2003). *L'Arbre des Familles*. Paris, Fr.: Éditions de La Martinière.
- En ce qui concerne la représentation des liens généalogiques, la maison ou le corps humain proposeront parfois leur structure, et çà et là les "rameaux" susciteront l'arbre, les feuilles et les fleurs. Entre le IXe et le Xlle siècle, on essaya toutes ces formules sans s'arrêter définitivement à aucune. On les combina même pour accumuler le plus grand nombre de sens possible. La clarification et la concentration sur l'image végétale survinrent à partir des Xlle-Xllle siècles, au terme d'une longue méditation des religieux et des maîtres des écoles, sur les valeurs symboliques attachées à la croissance et à la floraison ou à la fructification. Ici, l'arbre était imbattable. S'élançant vers le ciel, gros de son feuillage et de ses fleurs renaissant à chaque printemps, porteur de fruits sauvages, il nourrissait l'imagination, et sa métaphore constamment enrichie, abondamment commentée, fut reportée dans la représentation de la parenté, de la communauté des consanguins, de la généalogie. S'il s'agissait de restituer visuellement l'idée de fécondité, de descendance proliférante, de vigueur d'une souche à travers les âges, alors le végétal n'avait plus de rivaux. (Original version p. 37)
- Kosslyn, S. (1994). *Elements of Graph Design*. New York, NY: W.H. Freeman.
- Kosslyn, S. (2006). *Graph Design for the Eye and the Mind*. New York, NY: Oxford Univ. Press.
- Lakoff, G. & Johnson, M. (1980). *Metaphors We Live By*. Chicago, IL: Univ. Chicago Press.
- Lakoff, G. (1987) *Women, Fire, and Dangerous Things: What categories reveal about the mind*. Chicago, IL: Univ. Chicago Press.
- Meirelles, I. (2007). The Use of Metaphors in Dietary Visual Displays around the World. *Visible Language*, no. 41.3, 204-219, *Special Issue: Visual Metaphors in User Support*. Guest editors: Karel van der Waarde & Piet Westendorp.
- Montenegro, M. (2011). "Visualizing Human Development" at web-site *Visualizing.org* <<http://www.visualizing.org/visualizations/human-development-index-hdi-tree>>
- Murdoch, J. E. (1984). *Album of Science: Antiquity and the Middle Ages*. New York, NY: Charles Scribner's Sons.
- Pinker, S. (1990). A Theory of Graph Comprehension. In Roy Freedle (ed.). *Artificial Intelligence and the Future of Testing*. Hillsdale, NJ: Lawrence, pp. 73-126.
- Ware, C. (2004). *Information Visualization: Perception for Design Second Edition*. San Francisco, CA: Morgan Kaufmann.