

## How were the indicators of the year calculated?

In order to identify the indicators of the year, the following formula has been applied:

$$f(v, n, s) = 1 + \frac{v - \frac{s}{10}}{6} - \frac{1}{n}$$

with: v = average rating value, n = number of users, s = standard deviation

What does the formula do?

a) 1+

This is basically creating a basis which ensures that the result of the formula will never be lower than 0 and thus avoids miscalculations.

b)  $\frac{v - \frac{s}{10}}{6}$

It starts from taking the average rating value of an indicator and subtracts a possible standard deviation divided by 10. This is necessary because the deviation can be bigger than the actual average value and would then produce useless numbers but also it would give too much weight to the standard deviation. This result is then divided by the maximum value an indicator could have which is 6; this means effectively that we look at what percentage of the maximum an indicator achieves (after deducting the spread of results). Thus the bigger the average value and the smaller

the deviation, the closer the result of  $\frac{v - \frac{s}{10}}{6}$  is to 1.

Combined with a) this means that the better the rating value and the lower the standard deviation, the closer the total result approaches 2. The lower the value the closer it comes to 1.

c)  $\frac{1}{n}$

This creates a number which is closer to 0 the bigger the number of users of an indicator is. By subtracting this from the result of a) and b) we make sure that the number of users influences the result to an acceptable extend. The larger the number of users the more  $\frac{1}{n}$  has a limes of 0, i.e. the lesser the amount subtracted from a)+b). The smaller the number of users, the more  $\frac{1}{n}$  has a limes of 1, i.e. it uses up the basic value given by a).

d) Result

If the rating value is very small and only one person said so the overall value f will approach  $\frac{1}{6}$ . If the rating value is very high and many people used it with a high level of agreement, the f value will approach 2. In between, anything is possible but the standard deviation has the smallest of all effects

(which is ok because a high rating value with many users means that many people have to have rated it highly anyway). The number of users is very relevant and no large scale effects can take place.

Additionally, the formular will always produce a positive value regardless of number of users, rating value or standard deviation (it produces a Div#0 error for  $n=0$  so if noone used the indicator which is fine because those can then easily be filtered out).