

Matching methodology for WeCitizens' VAA

Establishing the candidate's profile

A candidate's profile is the result of all the answers to the questionnaire of the Vote Advice Application (VAA).

Both, the candidates and the voters have five possible answers : "I fully agree", "rather yes", "I don't know", "rather no", "I strongly disagree". This gradation is necessary to ensure that the candidates are able to differentiate themselves without infringing party discipline, if applicable.

Furthermore, voters may weigh the question: "important", "average importance" (default option) or "of low interest".

Should candidates wish to be considered for a vote recommendation, they will have to answer all the questions on the questionnaire. The voter, however, only has to answer the number of questions he/she wishes. Questions which have not been answered are not taken into consideration for the calculation.

Calculating the "match"

The "match" between the voter's profile and that of the candidate is based on the Euclidean space (geometric distance within a multidimensional space). In order to situate the answers within this space, they are given a numerical value according to the table below:

	Options given to answer each question		Value given
	to candidates	to voters	
I fully agree	x	x	100
rather yes	x	x	75
I don't know	x	x	50
rather no	x	x	25
I strongly disagree	x	x	0
no answer		x	

Thus, the Euclidean distance between a candidate and a voter is :

$$D = \sqrt{\sum_{i=1}^n (v_i - c_i)^2}$$

where : v_i = the voter's position with regard to question i
 c_i = candidate's position with regard to question i
 n = number of questions to which the voter has replied.

This distance is then correlated with the maximum distance:

$$D_{\max} = \sqrt{\sum_{i=1}^n (100)^2} = \sqrt{n \cdot 100^2}$$

resulting in the normalized distance : D / D_{\max}

A value indicating the level of match is achieved by subtracting the normalized distance of 1. In order to make this easier to read, the result is expressed in percentage points:

$$\text{Match} = 100 \cdot (1 - D/D_{\max})$$

The percentage point represents a measure of geometric correspondence. It does not refer to that part of the questionnaire to which the voter and the candidate provided the same answer.

This means of calculation provides a good balance between very simple approaches (e.g. one point for each concordant question, ...) and those which are more developed (e.g. factorial analysis, etc.).

Weighting answers

In any multi-criterial decision, it is advisable to weight the assessment criteria. This is why we have given the voter the possibility of weighting (w_i) each question differently:

2.5 = important

1 = normal weighting

0.4 = low importance

The preceding formula are then completed as follows:

$$D_w = \sqrt{\sum_{i=1}^n (w_i (v_i - c_i))^2}$$

$$D_{w,\max} = \sqrt{\sum_{i=1}^n (w_i \cdot 100)^2}$$

$$\text{Match}_w = 100 \cdot (1 - D_w/D_{w,\max})$$

where : v_i = the voter's position with regard to question i
 c_i = candidate's position with regard to question i
 w_i = voter's weighting with regard to question i
 n = number of questions to which the voter has replied.

Ranking

Match_w thus, calculated for each of the candidates which then allows them to be ranked in decreasing order of the value achieved for each one's Match_w .

When two or more candidates achieve the same score, they are ranked by taking into consideration their place on their respective electoral lists.