Online Job Advertisements

Sources Ranking Model

Francesco Trentini

Web
Intelligence
Hub

Landscaping OJA Web data sources

Deliverable D4.1 – Report with ranking model

Authors: Francesco Trentini

Approved by: Emilio Colombo, Mario Mezzanzanica

Version: 1

Date of Release: 2022-05-24

Contents

1. Ranking sources: model framework and goals

The landscaping activities produce a vast knowledge about Online Job Advertisements (OJAs) markets in Europe. This knowledge spans from the features and business model of OJA providers up to the market structure and the context in which they operate. Moreover, detailed information is registered on the characteristics of each portal as they appear online to jobseekers and prospective employers. The Final Landscaping Report organises this knowledge in a comprehensive way; nonetheless it does not give explicit operative directions for the subsequent phases of the analysis, which encompass identification of new sources to prioritise for agreements to access the data, ingestion and processing.

This role is fulfilled by the ranking model, which provides a framework to select the most relevant sources for development. The specificity of this evaluation framework is the involvement of many layers of information to cover different aspects of each source. Indeed, a qualitative assessment of sources' relevance in OJA market is performed, together with a quantitative assessment of adherence of each source to the desired characteristics that define the degree of quality of the data source to produce high-standard OJV and skill data in the current DPS. The final indicator is based on the combination of these two measures and allows to identify sources that are not eligible for being included in the final list of sources and to prioritise sources that should be engaged first through *ad-hoc* agreements with ownership to access the data.

The report discusses each of these steps as follows and integrates the description given in *D1.1 OJA landscaping methodological guide*. Section 2 introduces rationale and goals of the ranking model, and goes into details in the constituting elements: AHP score in section 2.1, ICEs' ranks in section 2.2 and the final score and the decision rule in sections 2.3 and 2.4 respectively. Section 3 presents an overview of the final ranking.

2.   The ranking model

The model that has been implemented by CRISP is constituted by two building blocks that are integrated into a final indicator that assesses the degree of priority of each source. The two building blocks are the AHP score and the ICEs' ranks.

The former is a quantitative measure of the quality of the sources. In this case quality is defined as the degree of desirability of the characteristics a source possesses. Desirability is itself defined according to the preferences that stakeholders have over the set of possible features of sources and how they compare to each other. The AHP model allows to compute, on the base of ratio scales in pairwise comparisons, a numerical score for each characteristic that represent a synthesis of the preferences of the stakeholders involved in the evaluation.

The latter, the ICE rank, is a qualitative measure of the importance of each source in the OJA market of the country where it is surveyed. The evaluation is based on known characteristics of the source in term of popularity in internet searches, stability in supplying high quality data and coverage of domains of occupation, regional and economic activity, integrated with ICEs personal knowledge of the OJA market in their country.

2.1.   The Analytic Hierarchy Process (AHP) score

The Analytic Hierarchy Process (AHP) is a method to support multi-criteria decision making, which allows as inputs pairwise comparisons of criteria. Criteria can be nested in different layers, which can also be subject to pairwise comparisons, providing, as a result, a level specific as well as global evaluation of priority of each criterion against all the others. The method combines preferences expressed by multiple stakeholders in a single output. An online tool assists in collecting the pairwise comparisons: https://bpmsg.com/ahp/ahp-hiergini.php?sc=AbezyP.

The criteria on which the method is applied is the relevant subset of all the possible source characteristics used to survey sources in activity *D2.1 Updated list of OJA sources, including their characteristics* and reported as Annex 1 in this document. Once the global priority is computed, the obtained values are mapped to the observed characteristics of each source and the algebraic sum of the values over these variables leads to the source specific AHP score.

2.2.   ICEs' ranks and comments

The role of ICEs is to evaluate the importance of each known source in terms of the relevance and specificity in the OJA market of their country of belonging. The operative activity requires to assign a ranking score and a comment on the specificities of each source. To inform the evaluation, CRISP has elaborated indicators three dimensions, popularity, stability and coverage.

The dimension of popularity captures the knowledge that jobseekers have of surveyed sources, by evaluating the volume of direct searches to a source. The computation is done on the performance of sources' relative interest as produced by Google Trends. The relative interest is computed with reference to the subset of Google Search queries belonging to the category of job listings. Therefore, the indicator can be interpreted as the relative popularity of each source among the searches directed to job listings. It is computed on 2021 data and is available for sources listed in Landscaping 2021 and 2017.

The second dimension is stability in the provision of OJAs data. The information is available only for sources already in the Data Production System (DPS). Each source–country pair is flagged as unstable if it does not meet at least one of a set of conditions, namely, on the maximum number of months without observations, on the variability of the time series and the number of outliers. Stability is enriched by additional information for each source, in cases in which it is not developed or currently not present in the final dataset.

Finally, for each source in the DPS, coverage indicators are calculated on data from 2019. Each source is assigned the count and relative frequency of vacancies computed for the source, broken down by ISCO08 Major groups, NUTS Macro-regions and NACE Macro-sectors. This computation is useful for ICEs to know the actual performance of known sources in providing OJA for the country[1] for different levels of detail, so that specificities can be assessed *ex ante*, as they emerge from the websites of the job portals and *ex post* after the ingestion and processing in the DPS.

2.3.   Final score

The AHP score and the ICEs' ranks are then elaborated to compute a final score. This measure provides the decision-making layer of information and, accordingly, must be simple without biasing or loosing information provided by input factors. Following this principle, AHP score and ICEs' ranks were mapped to the quartile of belonging in the respective distribution of values. Then five groups were defined to consider the joint distribution of AHP score and ICEs' ranks and mapped according to the scheme provided in Table 1. The lowest the score the highest the position in the joint distribution of AHP score and ICEs' ranks, meaning a higher priority assigned to the source.

Table 1 Score definition

Score	Definition	Cases (AHP score quartile, ICE rank quartile)
1	Sources with position in Q1 of ICE rank and Q1 of AHP score.	(Q1,Q1)
2	Sources with position in Q1 or Q2 of ICE rank and Q1 or Q2 of AHP score. Exclude the case (Q1,Q1).	(Q1,Q2),(Q2,Q1) and (Q2,Q2)
3	Sources with position in Q2 or Q3 of ICE rank and Q2 or Q3 of AHP score. Exclude the case (Q2,Q2).	(Q2,Q3),(Q3,Q2) and (Q3,Q3)
4	Sources with position in Q3 or Q4 of ICE rank and Q3 or Q4 of AHP score. Exclude the case (Q3,Q3).	(Q3,Q4),(Q4,Q3) and (Q4,Q4)
5	Sources with distance between position in ICE rank distribution and AHP score distribution larger than 1 quartile.	All the others, e.g. (Q1,Q4), (Q3,Q1)

---

[1] Please note that only a subset of the ads produced by each source has been considered in this count. Advertisements are filtered, keeping only OJAs that advertise a job in the country of interest, which are later deduplicated at the source level.

2.4. Decision rule

Once the measure has been set, a decision rule is needed to select the sources to be prioritised for inclusion in the DPS. Clearly the decision rule must take into account the fact that new countries participated in this Landscaping activity. The rule is therefore twofold. For countries already in the DPS, only new sources with scores 1 and 2 are candidates for development. New countries have an *ad-hoc* treatment. Switzerland, which registered a high number of sources, will have the same treatment as DPS countries. In Lichtenstein the threshold will be increased to 3, while in Iceland and Norway all potential sources will be candidates for development. The sources found in this latter group are prioritised in development to ensure faster convergence to the DPS standard.

3. Final ranking

The final ranking that results from the procedure presented here is exposed in *D4.2 - List of OJA sources for which data acquisition is propose* and presents the overall situation as well as the country specific sources ranking and selection. We report only the former in Table 2.

The prospect shows that 143 sources are expected to be developed, taking into account the fact that some sources are redundant as they are surveyed in new countries but already present in the DPS and will not require development.

Table 2 Count of sources listed in 2021 by country and score. Selected sources are highlighted – light blue for countries in the DPS, light green for new countries.

	Score					Total
	1	**2**	**3**	**4**	**5**	
AT	1				2	3
BE	2	1			9	12
BG		2	3	8	6	19
CH	7*	14	7	20	20	68
CY		2	2	1		5
CZ	1	1	6	3	6	17
DE		4	3	8	9	24
DK	1	3	4	4	2	14
EE	1		1	4	1	7
EL		2	1	4	2	9
ES	1	6	2	9	6	24
FI	1	1	4	3	4	13
FR	3	6		3	2	14
HR			1	5	1	7
HU	1	2	8	7	7	25
IE	3	1			5	9
IS	1	2	1	3	1	8
IT					5	5
LI	2§	5	6	3	9	25
LT	1	1	5	2	13	22
LU	1	4	5	4	11	25
LV		9	1	1	8	19
MT			2		1	3
NO	1	3		4	2	10
NL		4	3		4	11
PL	2	3	9	3	3	20
PT	2	4	4	4	1	15
RO		2	3	3	2	10
SE	1	3	9	4	9	26
SI	1	3	2	1	4	11
SK		1	3	5	5	14
UK		3	2	5	4	14
Total	34	92	97	121	164	508

Note: * Xing and Stellenanzeigen already in DPS in Germany; § Jobscout24.ch duplicate in CH; Xing.com already in DPS in Germany