Quality Assurance of
Content

neuvoo

The QA Evaluation

1. Every week, we pick a random sample of empcodes to assess the quality of their indexation on neuvoo. All the members of the Indexation team are assigned the same number of companies to review.

2. If only 60% of the data reviewed or less is accurate, the indexer will be penalized. If more than 3 penalizations are given within a 6 week period, it would mean the termination of the employment with neuvoo.

3. If a premium company is affected by an indexation error, this error will weigh twice as much as those affecting a non-premium company. This is why close attention is mandatory when editing a company that manages top content.

4. The QA Review will assess multiple elements, each one representing a value assigned according to the importance of the error found.

Indexation Errors Specification

In this section, you will find all the different types of errors indexers can bump into when adding new content on neuvoo. Each error costs a certain amount of points that will be deducted from the total of 100 points. A score of 100 means the spider doesn't have any of the errors described below.

Note: Errors with * are classified as critical which means they will automatically be valued as 0 because they are too significant and we can't let them happen.

Expected Jobs (code 1100 - 15 points): We were not able to extract all jobs from one of the job sites we are indexing for one (or more) of the following reasons:

- 1101*: The pagination is not coded.
- 1102: The pagination runs to a certain point and then fails.
- 1103: The spider is not extracting all the jobs because some categories are missing.
- 1104: The spider is not extracting all the jobs because some languages are missing.
- 1105: The spider is not extracting all the jobs because some multi locations are missing
- 1106: The spider is not extracting all the jobs.

Data Extraction (code 1200 - 20 points): Failed data extraction due to an issue with the format or content of one or several fields that need to be extracted to complete the indexation process.

- 1201: The empcode has capital letters or any special character different from "-" (the only special character allowed is "-").
- 1202: *Empname*, *Title* or *Location include extra* or irrelevant information.
- 1203*: *Source_empname* is not extracted and it should be.
- 1204*: *The Empname, Title, Location, URL or Application Ema*il was not extracted from the job site.
- 1205: *The Salary or Job Type* was not extracted from the job site.
- 1206*: The *Empname* does not match the company indexed or is misspelled.
- 1207: The homepage does not match the company indexed.
- 1208*: The *Empname*, *Title, Location*, *URL*, *Salary*, *URL Tracker*, *Job Type* or *Application Email,* extracted is not the correct one. (Wrong selector).
- 1209: Some information regarding the location (City, Region, Country) was not extracted from the original job site.
- 1210: The Job Description includes the contact information of the employer or hiring manager.
- 1211: The Job Description includes information not related to the actual position.
- 1212: The Job Description is incomplete.
- 1213: The Job Description is missing.
- 1214: The Job URL redirects to a list of jobs and not the specific listing.
- 1215: Dev Comment missing when specific conditions are requested.
- 1216*: The *CPC value,* extracted is not the correct one. (Wrong selector).

Valid Content (code 1300 - 20 points): Job doesn't meet our quality requirements to be valid.

- 1301: Expired jobs.
- 1302: The closing date or posted date was not coded in the spider when available on the job site.
- 1303*: Duplicated content.
- 1304: Companies flagged as a scam.
- 1305*: The job site indexed belongs to a job aggregator.
- 1306*: The job site indexed belongs to a fake job board.
- 1307*: The content indexed is not a job. It is another type of content.
- 1308*: Inappropriate content (related to drugs, prostitution or illegal activities).
- 1309*: Specific conditions missing or not properly coded.

Company Classification (code 1400 - 15 points): The source to the website where the job was originally published was not correctly categorized (company, placement agency, job board, etc).

- 1401: Company classified as a job board.
- 1402: Company classified as a placement agency.
- 1403*: Job board classified as a company.
- 1404: Job board classified as a placement agency.
- 1405*: Placement agency classified as a company.
- 1406: Placement agency classified as a job board.

Spider Performance (code 1500 - 15 points): Spider not running properly.

- 1501*: The spider is reindexing the jobs in each run.
- 1502: The spider is extracting 0 jobs even when jobs are available.
- 1503: The system detects that the runs are not normal (Boo Error).
- 1504: The spider is stuck in a never-ending process (Stuck Error).

Logo (code 1600 - 15 points): Indexation of the wrong company logo.

- 1601*: Default logo is different from the one displayed on the original job site.
- 1602: Logo is not 100x100 and looks pixelated.
- 1603: Logo was successfully extracted from the XML feed but there is no coded version of it.

- 1604: Logo successfully extracted from the XML feed but not coded properly.
- 1605: Source empname logo is different from the one displayed on the original job site.