

## Evalucacion QA

Cada semana, seleccionamos una muestra aleatoria de empcodes para evaluar la calidad de su indexación en neuvo. A todos los miembros del equipo de Indexación se les asigna el mismo número de empresas para revisar.

Si solo el 60% de los datos revisados o menos son precisos, el indexador será penalizado. Si se dan más **de 3 penalizaciones en un período de 6 semanas**, significaría la terminación del empleo con neuvo.

si una **empresa premium** se ve afectada por un error de indexación, este error pesará el doble que los que afecten a una **empresa no premium**. Esta es la razón por la cual es obligatorio prestar mucha atención cuando se edita una empresa que administra contenido superior.

La Revisión de QA evaluará múltiples elementos, cada uno representando un valor asignado de acuerdo a la importancia del error encontrado.

### Especificaciones de errores de indexacion

En esta sección, encontrará todos los diferentes tipos de errores con los que se pueden encontrar los indexadores al agregar contenido nuevo en neuvo. **Cada error cuesta una cierta cantidad de puntos que se descontarán del total de 100 puntos.** Una puntuación de 100 significa que el spider no tiene ninguno de los errores que se describen a continuación.

**Nota: Los errores con \* se clasifican como críticos, lo que significa que se valorarán automáticamente como 0 porque son demasiado importantes y no podemos permitir que sucedan.**

Trabajos esperados (**código 1100 - 15 puntos**): No pudimos extraer todos los trabajos de uno de los Job site que estamos indexando por una (o más) de las siguientes razones

- 1101\*: La paginación no está codificada.
- 1102: la paginación se ejecuta hasta cierto punto y luego falla.
- 1103: La araña no está extrayendo todos los trabajos porque faltan algunas categorías.
- 1104: La araña no está extrayendo todos los trabajos porque faltan algunos idiomas.
- 1105: la araña no está extrayendo todos los trabajos porque faltan algunas ubicaciones múltiples
- 1106: La araña no está extrayendo todos los trabajos.

Extracción de datos (**código 1200 - 20 puntos**): extracción de datos fallida debido a un problema con el formato o contenido de uno o varios campos que deben

extraerse para completar el proceso de indexación.

- 1201: el empcode tiene letras mayúsculas o cualquier carácter especial diferente de "-" (el único carácter especial permitido es "-").
- 1202: Empname, Title o Location incluyen información extra o irrelevante.
- 1203\*: Source\_empname no se extrae y debería hacerlo.
- 1204\*: el Empname, el cargo, la ubicación, la URL o el correo electrónico de la aplicación no se extrajeron del sitio de trabajo.
- 1205: el salario o el tipo de trabajo no se extrajo del sitio de trabajo.
- 1206\*: El Empname no coincide con la empresa indexada o está mal escrito.
- 1207: La página de inicio no coincide con la empresa indexada.
- 1208\*: el Empname, el Title, la Location, la URL, el Salary, el URL Tracker, el tipo de trabajo o el correo electrónico de la aplicación extraídos no son los correctos. (Selector incorrecto).
- 1209: parte de la información sobre la Location(city, Region, Country) no se extrajo del sitio de trabajo original.
- 1210: La Job Description incluye la información de contacto del empleador o gerente de contratación.
- 1211: La Job Description incluye información no relacionada con el puesto real.
- 1212: La Job Description está incompleta.
- 1213: Falta la Job Description .
- 1214: la Job URL redirige a una lista de trabajos y no a la lista específica.
- 1215: Falta el Dev Comment cuando se solicitan condiciones específicas.
- 1216\*: El valor de CPC extraído no es el correcto. (Selector incorrecto).

Contenido válido (**código 1300 - 20 puntos**): el trabajo no cumple con nuestros requisitos de calidad para ser válido.

- 1301: Jobs caducados.
- 1302: La fecha de cierre o la fecha publicada no se codificó en el Spider, cuando esta disponible en el sitio de trabajo.
- 1303\*: contenido duplicado.
- 1304: Companies marcadas como estafa.
- 1305\*: El job site indexado pertenece a un agregador de trabajo.
- 1306\*: El job site indexado pertenece a una bolsa de trabajo falsa.
- 1307\*: El contenido indexado no es un trabajo. Es otro tipo de contenido.
- 1308\*: Contenido inapropiado (relacionado con drogas, prostitución o actividades ilegales).
- 1309\*: Condiciones específicas faltantes o codificadas incorrectamente.

Clasificación de la empresa (**código 1400 - 15 puntos**): La fuente del sitio web donde se publicó originalmente el trabajo no se categorizó correctamente (company,

placement agency, job board, etc.).

- 1401: Company clasificada como una Jobboard(bolsa de empleo).
- 1402: Company clasificada como una Placement agency
- 1403\*: Jobboard clasificada como Company.
- 1404: Jobboard clasificada como Placement agency.
- 1405\*: Placement agency clasificada como Company.
- 1406: Placement agency clasificada como Jobboard.

Spider Performance (**código 1500 - 15 puntos**): Spider no funciona correctamente.

- 1501\*: The spider esta reindexando los trabajos en cada ejecución.
- 1502: The spider esta extrayendo 0 Jobs incluso cuando hay Jobs disponibles.
- 1503: The Sistema detecta que las ejecuciones no son normales(Boo Error).
- 1504: The spider esta atascada en un proceso interminable (Stuck Error).

Logotipo (**código 1600 - 15 puntos**): Indexación del logotipo de la Company incorrecto.

- 1601\*: el logotipo predeterminado es diferente del que se muestra en el Job site original.
- 1602: el logotipo no es de 100x100 y se ve pixelado.
- 1603: el logotipo se extrajo correctamente de la fuente XML, pero no existe una versión codificada del mismo.
- 1604: el logotipo se extrajo correctamente de la fuente XML pero no se codificó correctamente.
- 1605: el logotipo del Source Emprname es diferente del que se muestra en el Job site original.