



CONGRESO INTERNACIONAL
RED UNIVERSIDAD
EMPRESA ALCUE

EL USO DE LA MINERÍA DE DATOS PARA LA OBTENCIÓN DE INFORMACIÓN DEL MERCADO LABORAL.

Autor: Ing. Jesús Luyo

OBSERVATORIO LABORAL

5 líneas de investigación:

- Panorama económico laboral
- Estudios sectoriales
- Estudios por áreas funcionales
- Investigaciones de demanda laboral en profesiones vigentes
- Investigaciones sobre demanda laboral en nuevas profesiones



INFORMACIÓN REQUERIDA – DEMANDA LABORAL

Perspectivas y situación actual

- Especializaciones requeridas
- Conocimientos requeridos
- Sectores que demandan profesionales
- Información cruzada





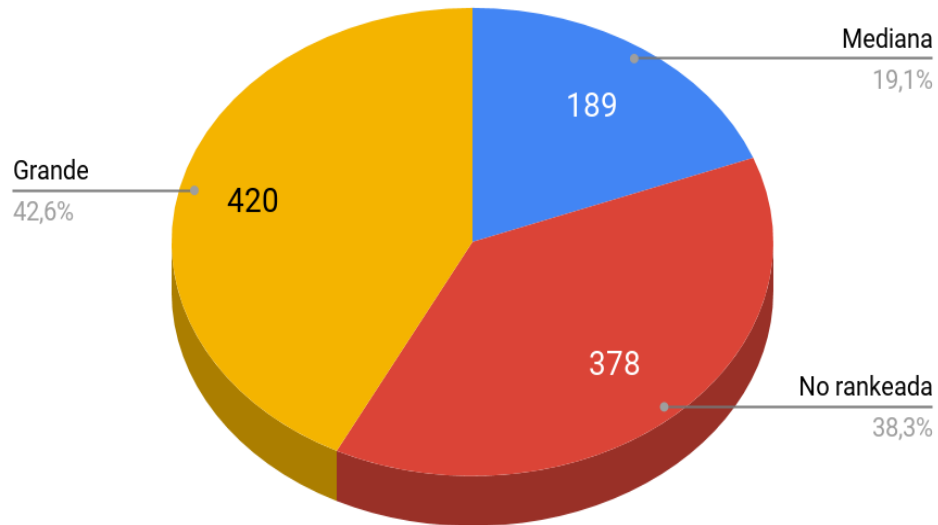
METODOLOGÍAS PARA RECOPIRAR INFORMACIÓN

- Tradicional
- Con aplicaciones de tecnología de información

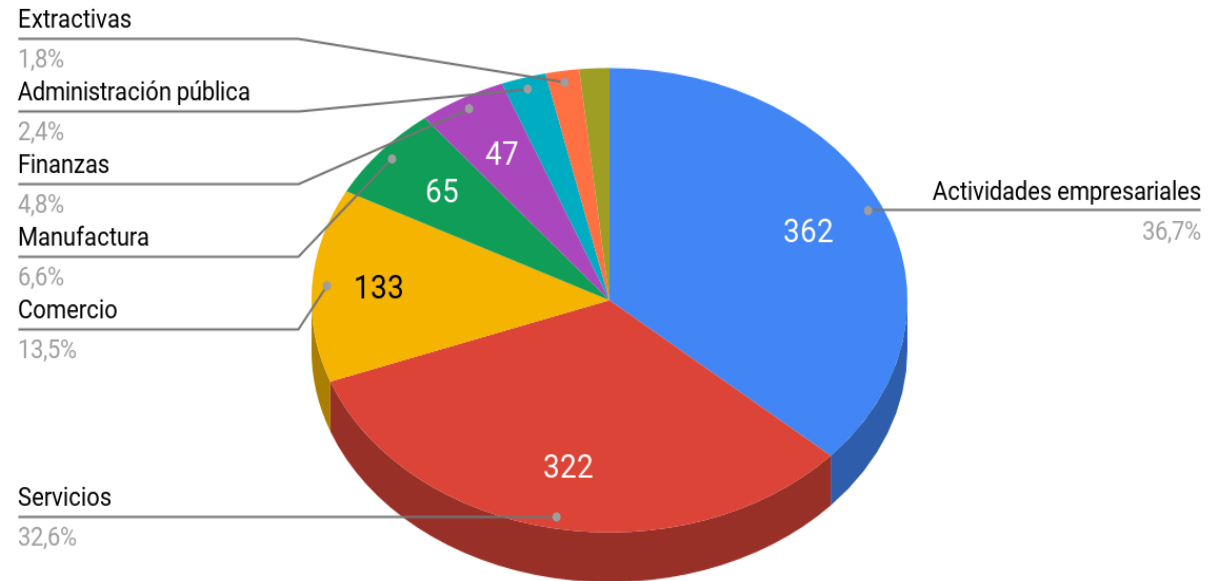
UN EJERCICIO CON METODOLOGIAS TIC

DEMANDA LABORAL PROFESION 1

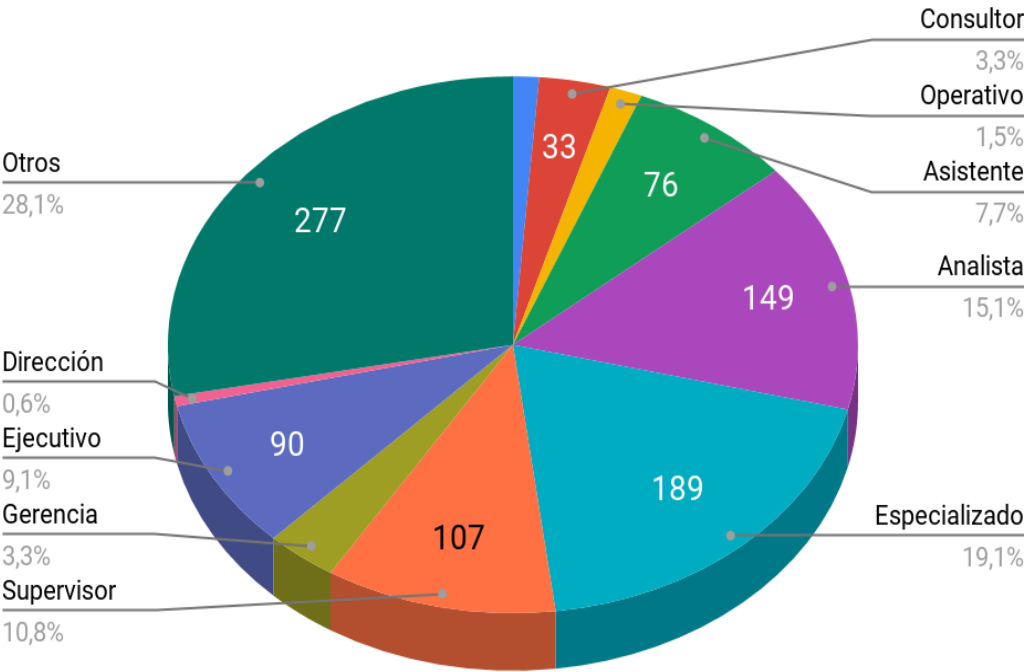
Según tamaño de empresa



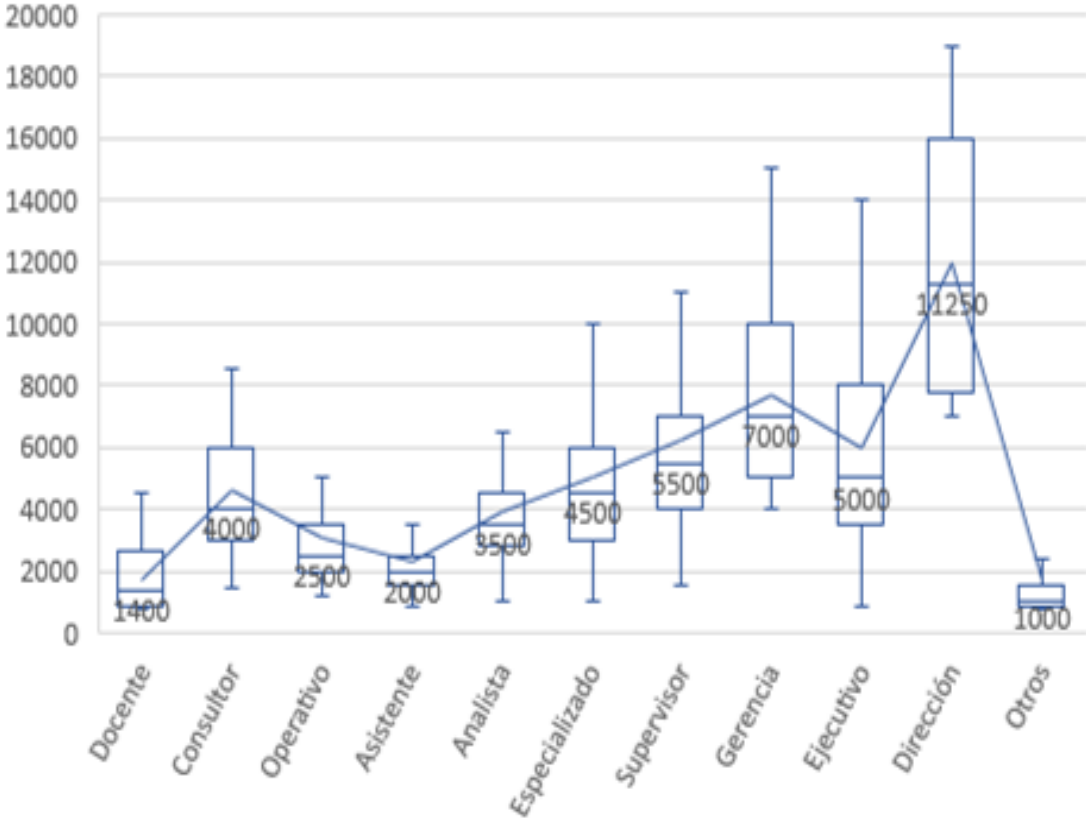
Según sectores de actividad



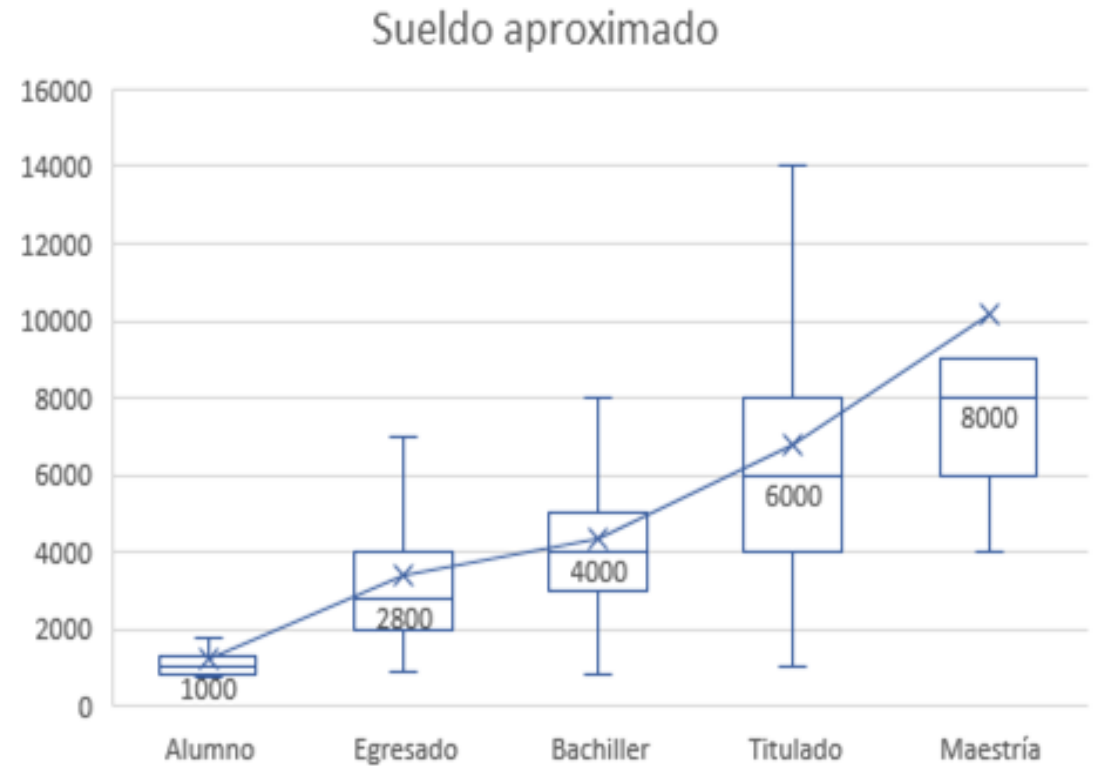
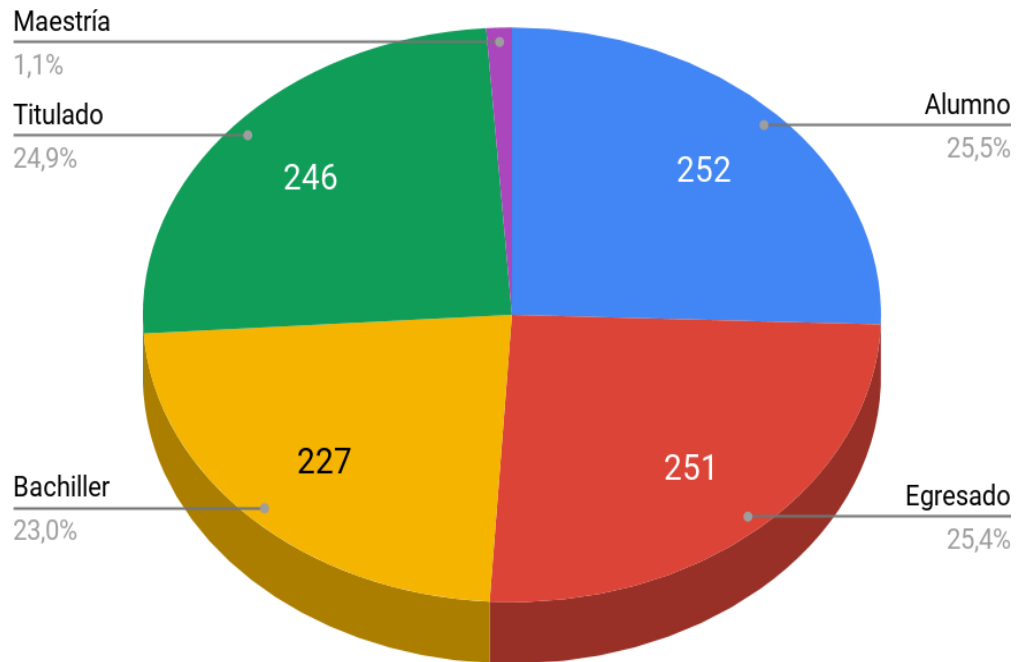
SEGÚN CARGO



Sueldo aproximado



SEGÚN NIVEL DE ESTUDIOS DEMANDADOS



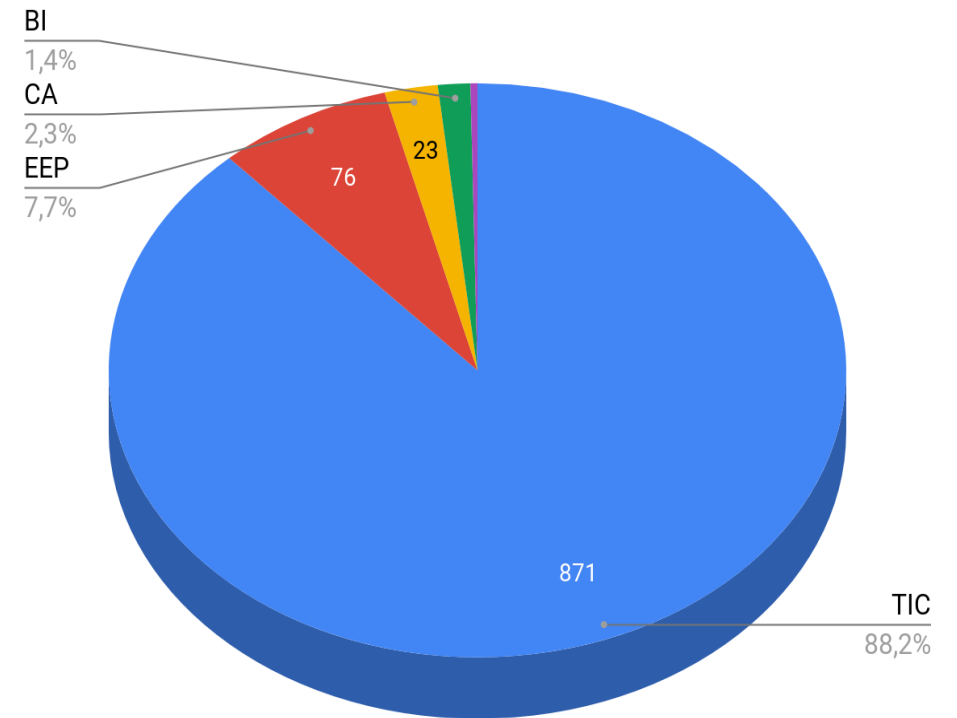
CONOCIMIENTOS ADICIONALES

Software	# conv
Ofimática	707
Programación	330
Bases de datos	159
ERP	94
Data warehousing	82
Software de gestión	66
CAD	59
Minería de textos	50
Seguridad informática	27

Idioma	# conv
Inglés	645
Portugués	4

PERFIL PROFESIONAL

Área	# conv
Tecnologías de Información y Comunicación (TIC)	871
Electricidad y Electrónica de Potencia (EEP)	76
Control y Automatización (CA)	23
Bioingeniería (BI)	14
Circuitos y Sistemas (CS)	3



PERFIL PROFESIONAL: CRUCES

Área	Tecnologías de Información y Comunicación	Electricidad y Electrónica de Potencia	Control y Automatización	Bioingeniería	Circuitos y Sistemas
Tecnologías de Información y Comunicación	871	18	52	2	0
Electricidad y Electrónica de Potencia	8	76	14	3	0
Control y Automatización	0	1	23	0	0
Bioingeniería	0	0	0	14	0
Circuitos y Sistemas	0	0	0	0	3

EL INFORME EN CIFRAS

El L4 es un informe cuantitativo de análisis de la demanda laboral para una determinada carrera.

Fuentes de Información:

- BT PUCP
- Fuente Externa 1
- Fuente Externa 2

INFORME L4

Campos de análisis:

- Áreas de Conocimiento
- Cargos
- Estudios
- Idiomas
- Software
- Sectores y Tamaño de Empresa



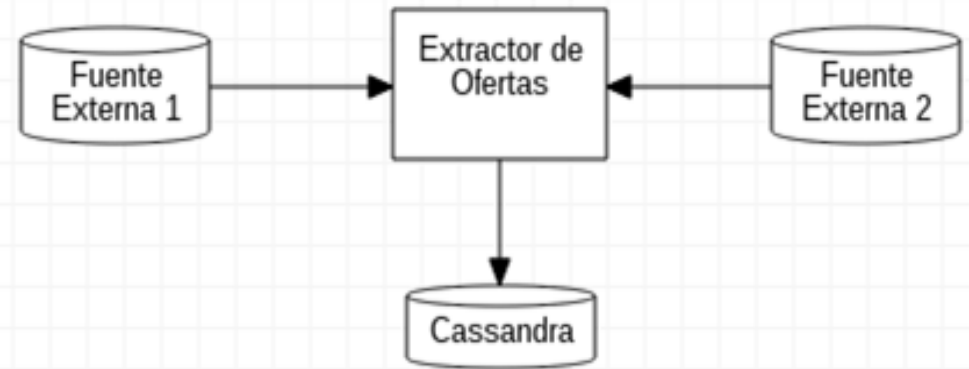
PASOS

1. Extracción de Convocatorias
2. Generación de Diccionarios
3. Procesamiento de Convocatorias
4. Generación de Reportes



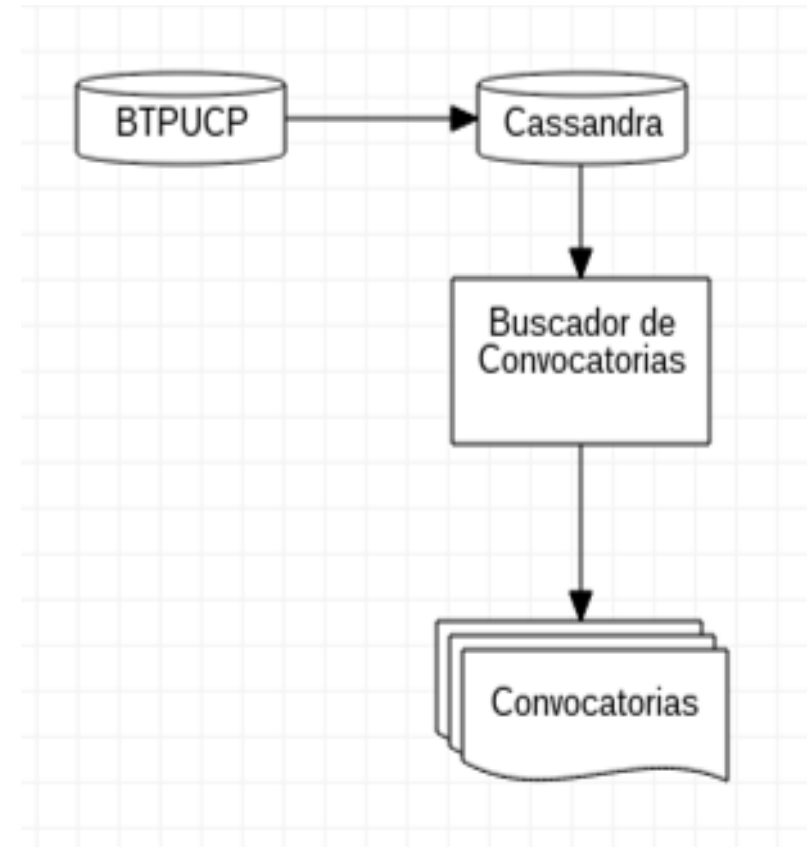
EXTRACCIÓN DE OFERTAS

- Se obtienen las ofertas de dos fuentes externas.
- El programa guarda las convocatorias extraídas a una nueva base de datos (Cassandra)



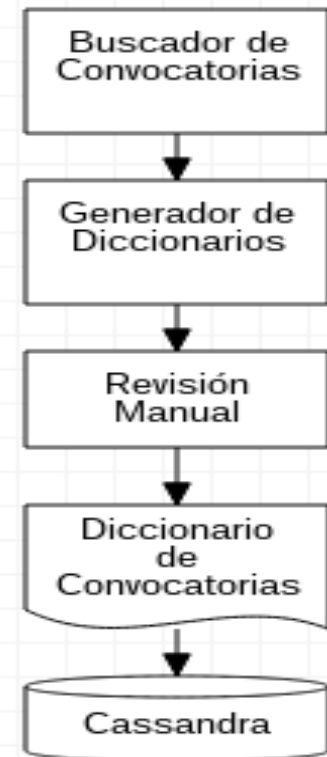
BUSCADOR DE CONVOCATORIAS

- Se incluyen las convocatorias de la BTPUCP
- Se buscan convocatorias por:
- Fuente (BTPUCP o cualquiera de las fuentes externas)
- Rango de Fechas
- Carrera



GENERADOR DE DICCIONARIOS

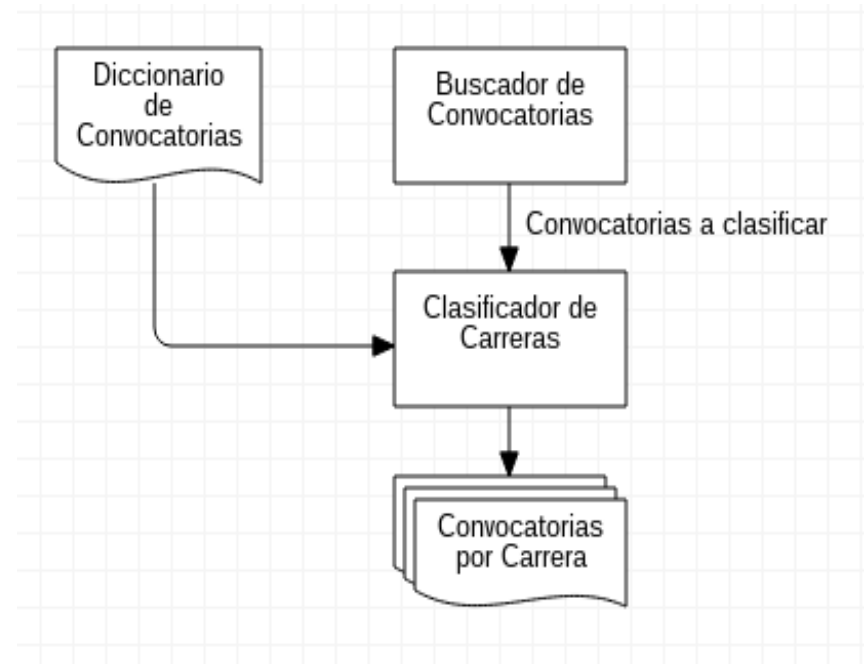
- Se obtiene un listado de las palabras más representativas de un conjunto de convocatorias
- El listado de palabras son revisadas manualmente y guardadas en la base de datos.



PROCESAMIENTO DE CONVOCATORIAS

Clasificador de carreras

- Para las fuentes externas, las convocatorias no cuentan con las carreras a la que están dirigidas.
- El clasificador de carreras determina si una convocatoria está dirigida o no a una determinada carrera.

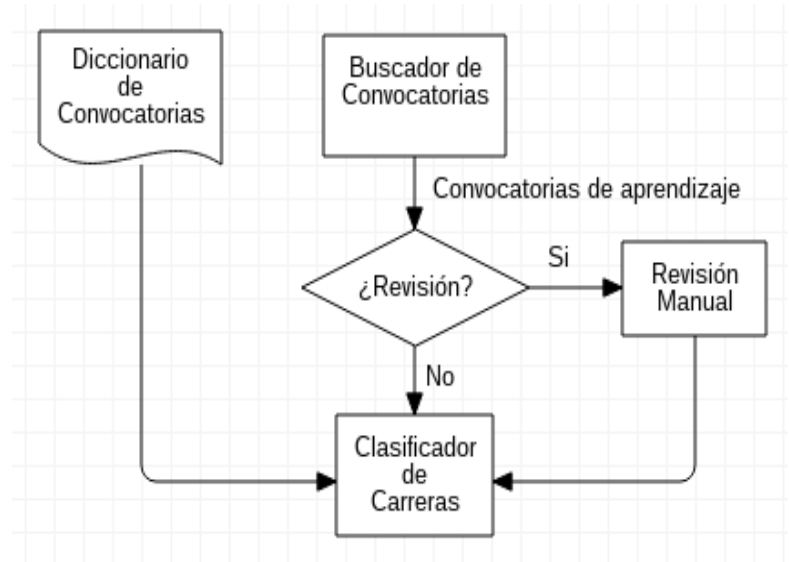


PROCESAMIENTO DE CONVOCATORIAS

Clasificador de Carreras

Los requisitos del clasificador de Carreras son:

- Conjunto de Convocatorias de aprendizaje
- (BTPUCP o revisadas manualmente)
- Diccionario de la carrera



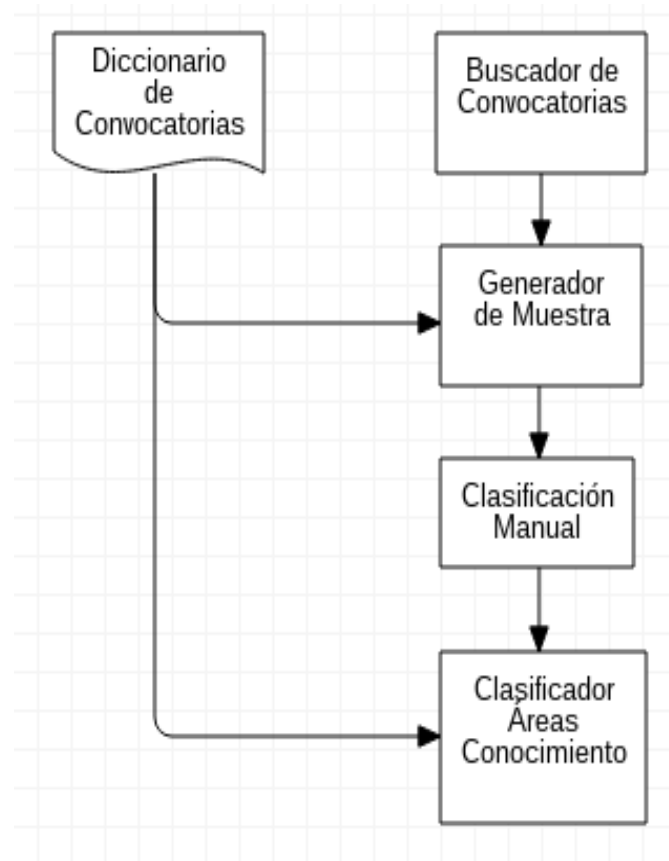
PROCESAMIENTO DE CONVOCATORIAS

Clasificador de Áreas de Conocimiento

Obtiene las áreas de conocimiento a las que va dirigida una convocatoria

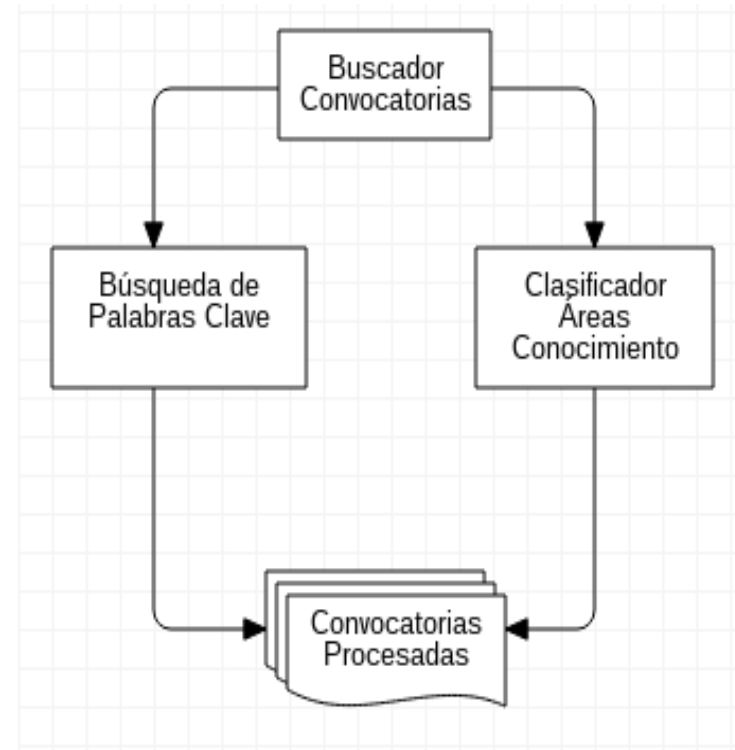
Los requisitos del clasificador de áreas de conocimiento son:

- Muestra de convocatorias etiquetadas manualmente
- Diccionario de la carrera



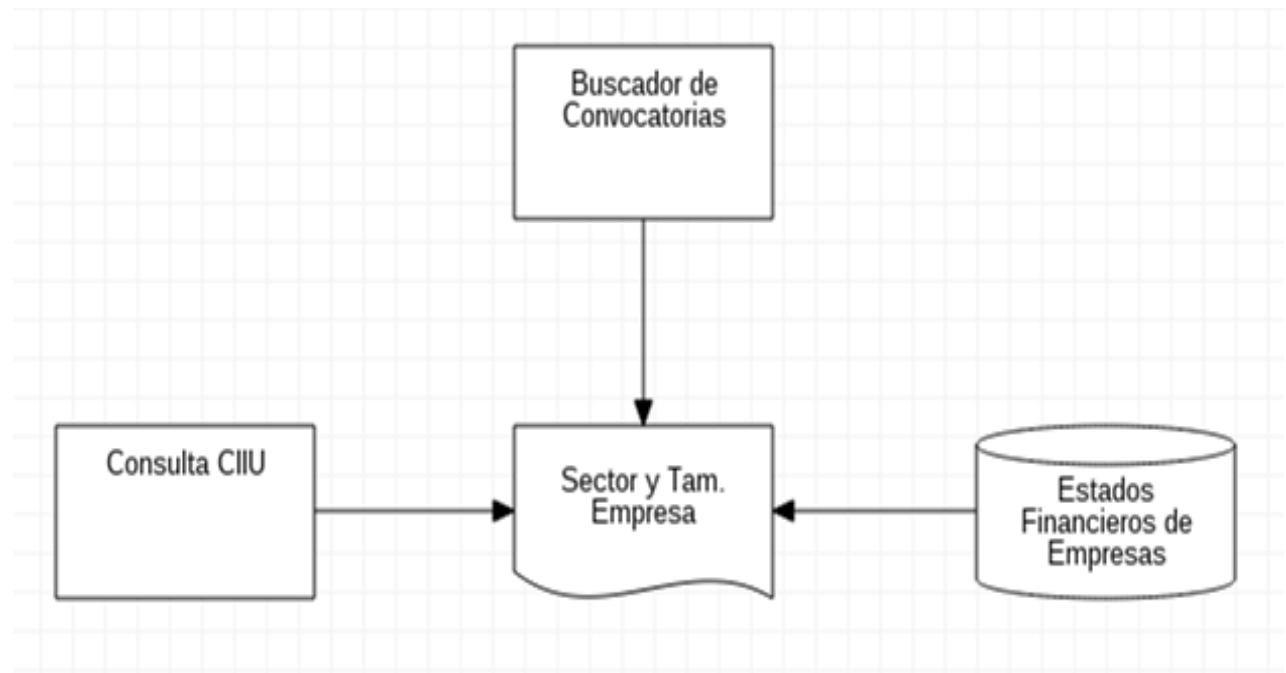
PROCESAMIENTO DE CONVOCATORIAS

- Para los otros campos de análisis, se sigue un proceso que consiste en la búsqueda de palabras clave.



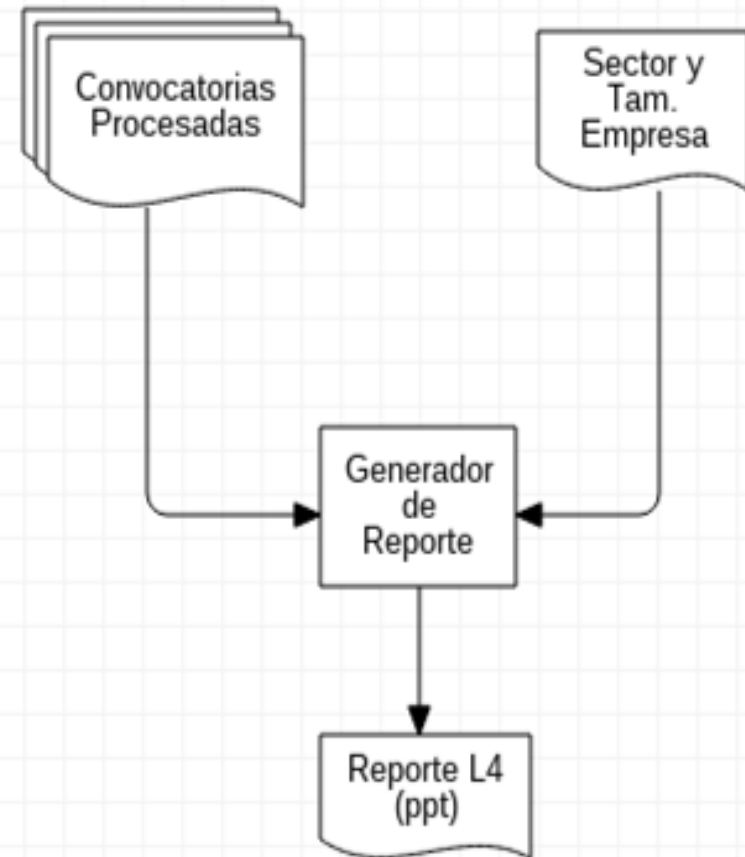
PROCESAMIENTO DE CONVOCATORIAS

- Un análisis adicional sobre las convocatorias de la BTPUCP se realiza para obtener el sector y tamaño de las empresas.



GENERACIÓN DE REPORTE

- Con las convocatorias procesadas y el análisis de sectores y tamaño de empresas, se genera un reporte L4 completo.





LO APRENDIDO

- No está mal empezar en pequeño, se fortalecen los conceptos.
- En páginas públicas las estructuras no son para investigar
- Hay mucha data con errores en páginas públicas.
- Utilizar software libre
- Ayuda mucho una persona con una visión general de recursos disponibles
- En las clasificaciones humanas pueden haber sesgos muy fuertes.



LO APRENDIDO

- Minuciosa revisión de los fundamentos teóricos de las aplicaciones utilizadas.
- Hay que planificar la oferta del servicio para que tenga aceptación.
- Es conveniente reducir el esfuerzo de unidades internas.
- Siempre aparecen nuevas posibilidades



¡GRACIAS!

jluyo@pucp.edu.pe