Proyecto Integra Job – Plataforma inclusiva de empleo

Documento de arquitectura de software

**Versión:** 1.0

**Fecha:** Octubre 2025

**Autores:** Carlos Colmenarez, Francisco Eljuri y Javier Rivas

**Responsable:** Javier Rivas

# Índice:

[Índice: 2](#_xin66y66okcr)

[**Introducción: 3**](#_orsnnjubc4pu)

[Propósito: 3](#_yw48cv9xaxr6)

[Alcance: 3](#_7drh4r62k5gr)

[Audiencia: 3](#_kbffa6ykbmar)

[**Control de versiones: 3**](#_799ko76b85dx)

[**Visión general del sistema: 4**](#_489zjzidxgqq)

[Metas de arquitectura: 4](#_26gs5npjpmst)

[Componentes principales: 5](#_gkrm3zxnfmg2)

[**Patrón arquitectónico: 6**](#_s58r8oue3b65)

[Flujo de integración con la API BNE: 6](#_myovnn5u76a9)

[**Modelo de datos: 8**](#_9066971m450f)

[Entidad principal: OfertaJob  
Relaciones principales: 8](#_jcinrwfpi9ax)

[Tablas destacadas: 8](#_a5ib42alvnwi)

[**Integraciones externas: 9**](#_6cmqn59hiyv3)

[API bolsa nacional de empleo (BNE) 9](#_r6azab6pnasm)

[Otros posibles orígenes: 9](#_9uj23xhgthbo)

[**Seguridad: 10**](#_1z3otqf6r7c3)

[**Calidad del software: 10**](#_sgcok9fg0iw5)

[**Estrategia de despliegue y mantenimiento: 10**](#_hz69aaav44tr)

[Despliegue local para desarrollo: 10](#_xxebvv6mddsd)

[Despliegue producción (planificado): 11](#_8vsjy9tmow1s)

[**Glosario: 11**](#_lhnble7d8bd6)

[**Referencias: 11**](#_irkcn9o2x7ya)

# Introducción:

## Propósito:

Este Documento de Arquitectura de Software (DAS) define la estructura técnica, las decisiones de diseño y los lineamientos que orientan el desarrollo, despliegue y mantenimiento del sistema **IntegraJob**.

Su propósito es asegurar que todos los componentes tecnológicos trabajen de manera coherente, escalable y segura, en conformidad con los objetivos del proyecto Capstone.

## Alcance:

IntegraJob es una **plataforma inclusiva de intermediación laboral** que centraliza ofertas de empleo provenientes de diversas fuentes, especialmente la **API de la Bolsa Nacional de Empleo (BNE)**, y las adapta para personas con Neurodivergente o con necesidades de accesibilidad.

El sistema incluye módulos de integración (ETL Python Analytics), backend (Spring Boot), frontend (React) y bases de datos (MySQL).

## Audiencia:

Equipo de desarrollo (backend, frontend y ETL)

QA y Testing Automation

Stakeholders y supervisores académicos

Clientes o usuarios interesados en la interoperabilidad del sistema

# Control de versiones:

| **Versión** | **Fecha** | **Autor** | **Descripción del cambio** |
| --- | --- | --- | --- |
| 1.0 | 17 oct 2025 | Javier Rivas | Primera versión completa del DAS |

# Visión general del sistema:

IntegraJob es una **plataforma modular**, basada en microservicios, que recopila y clasifica ofertas laborales en función de su accesibilidad, características y ubicación, para conectarlas con usuarios y empresas inclusivas.  
 Su meta es **reducir las barreras de acceso al empleo** mediante automatización, integración de fuentes externas y una interfaz web accesible.

## Metas de arquitectura:

* Interoperabilidad (entre múltiples fuentes de datos laborales)
* Escalabilidad horizontal (Docker + servicios independientes)
* Mantenibilidad (arquitectura por capas y documentación técnica)
* Seguridad (OAuth2 y roles por entidad)
* Accesibilidad y experiencia de usuario inclusiva

Arquitectura general:

**Tipo:** Arquitectura en microservicios contenedorizada (Docker)  
**Patrón:** MVC (Modelo–Vista–Controlador)  
**Estilo de comunicación:** RESTful API + HTTP/JSON

## Componentes principales:

| Componente | Descripción funcional | Tecnologías / Infraestructura |
| --- | --- | --- |
| Frontend (IntegraJob React) | Aplicación web accesible que permite la interacción de los usuarios finales con el sistema. Gestiona búsquedas, filtros y perfiles laborales. | React.js, Axios, Tailwind, Node 18, Docker |
| Backend (IntegraJob API) | Servicio central desarrollado en Spring Boot que gestiona las entidades principales (usuarios, empresas, ofertas, taxonomía). Expone una API REST. | Java 17, Spring Boot 3, JPA (Hibernate), Lombok, MySQL 8, Docker |
| Python Analytics (ETL – Data Integration Layer) | Módulo encargado de extraer, transformar y cargar datos desde fuentes externas (Computrabajo, LinkedIn). Normaliza las ofertas según la taxonomía de IntegraJob y publica datos hacia la base de datos central o API REST. | **Lenguaje:** Python 3.11 **Framework:** Django REST Framework **Base de datos:** PostgreSQL 15 **Servidor:** Gunicorn + Nginx **Infraestructura:** Docker Compose (entornos local y prod), volúmenes persistentes (pgdata, static, media) |
| Base de Datos Central (IntegraJobDB) | Motor de persistencia principal utilizado por el backend de Spring Boot para almacenar la información normalizada de las ofertas, usuarios y taxonomía. | MySQL 8, Docker Volume mysql\_data |
| Orquestador de Contenedores | Administra los servicios del ecosistema en entornos de desarrollo y despliegue. | Docker Compose v3.9 |
| Control de versiones y CI/CD | Control de cambios, ramas de integración y despliegue automatizado entre entornos. | GitHub (máquinas-corp), Git Flow, GitHub Actions (planificado) |

# Patrón arquitectónico:

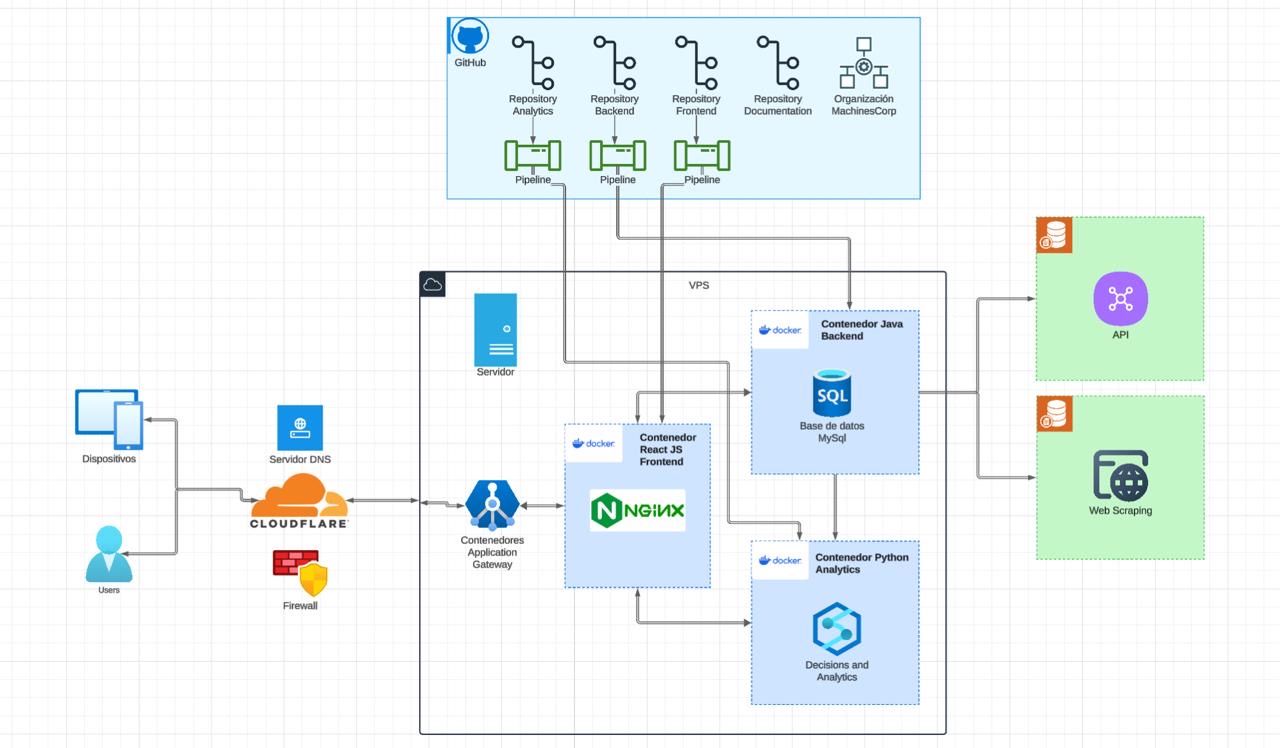
**MVC (Multi-Layer Model):** separa presentación, lógica y persistencia.

**Microservicios:** cada módulo independiente se comunica vía REST.

**ETL asíncrono:** Python Analytics opera fuera del backend, sincronizando vía API o exportación de datos.

## Flujo de integración con la API BNE:

1. *BNETokenService* (SB) solicita token OAuth2 (Client Credentials Grant).
2. *BneJobService* consume */JobOfferingsService/v1/1.0.0/jobofferings/active.*
3. Python Analytics recibe el payload, normaliza los campos y clasifica según taxonomía.
4. Los datos se persisten en MySQL.



# Modelo de datos:

Se anexa documento con script de la base de datos.  
 (script-integrajobdb-202510170147.sql)

### **Entidad principal:** OfertaJob Relaciones principales:

* OfertaJob → Empresa (N:1)
* OfertaJob → Ubicacion (N:1)
* OfertaJob → Caracteristica (N:M mediante OfertaCaracteristica)
* Usuario → Acteco (N:M)

### Tablas destacadas:

* empresa
* oferta\_job
* ubicacion
* caracteristica, agrupacion, sinonimo
* oferta\_caracteristica (relación cruzada)
* usuario, acteco

# Integraciones externas:

## API bolsa nacional de empleo (BNE)

* Autenticación: OAuth2 (Bearer Token)
* Endpoint principal: /JobOfferingsService/v1/1.0.0/jobofferings/active
* Parámetros clave: limit, offset, q, jobLocationIdentifier
* Límite: 200 solicitudes por minuto
* Entorno de pruebas: https://test.api.bne.cl

## Otros posibles orígenes:

* Computrabajo (web scraping)
* LinkedIn Jobs API
* ChileValora (certificaciones)

# Seguridad:

* Cifrado de credenciales mediante application.yml (Spring Boot).
* Tokens almacenados localmente con fecha de expiración (expires\_at).
* Control de acceso por roles (Usuario, Empresa, Administrador).
* Filtrado CORS habilitado solo para dominios autorizados.

# Calidad del software:

| **Actividad** | **Herramienta** | **Propósito** |
| --- | --- | --- |
| Pruebas unitarias | JUnit / pytest | Validar métodos individuales y ETL Python |
| Pruebas de integración | Postman | Validar flujo API ↔ BNE |
| Reportes QA | Excel | Documentar casos de prueba y resultados |
| Control de versiones | Git Flow | Rama Production ↔ staging ↔ feature |

# Estrategia de despliegue y mantenimiento:

## Despliegue local para desarrollo:

* **Docker Compose local** para Python Analytics (autorreload Django).
* **Spring Boot local + MySQL Dockerizado.**
* **React en modo desarrollo (3000).**

## Despliegue producción (planificado):

* Backend en **AWS EC2**
* Base de datos **AWS RDS**
* Python Analytics + Nginx en contenedor independiente
* Integración CI/CD con GitHub Actions
* Backups diarios y rollbacks controlados

## 

# Glosario:

| **Término** | **Definición** |
| --- | --- |
| **ETL** | Proceso de Extracción, Transformación y Carga de datos. |
| **BNE** | Bolsa Nacional de Empleo de Chile. |
| **API REST** | Interfaz basada en protocolo HTTP y formato JSON. |
| **Taxonomía Integra Job** | Estructura de clasificación de características de empleo por agrupación y sinónimos. |
| **Docker Compose** | Herramienta para definir y ejecutar aplicaciones multicontenedor. |

# Referencias:

**Guía Técnica de Integración API BNE**, Subsecretaría del Trabajo, Versión 03 (2025).

**Guía Estudiante – Fase 2 Desarrollo Proyecto APT**, Duoc UC (2025).

Documentación propia de repositorios *datapi* y *python-analytics*.