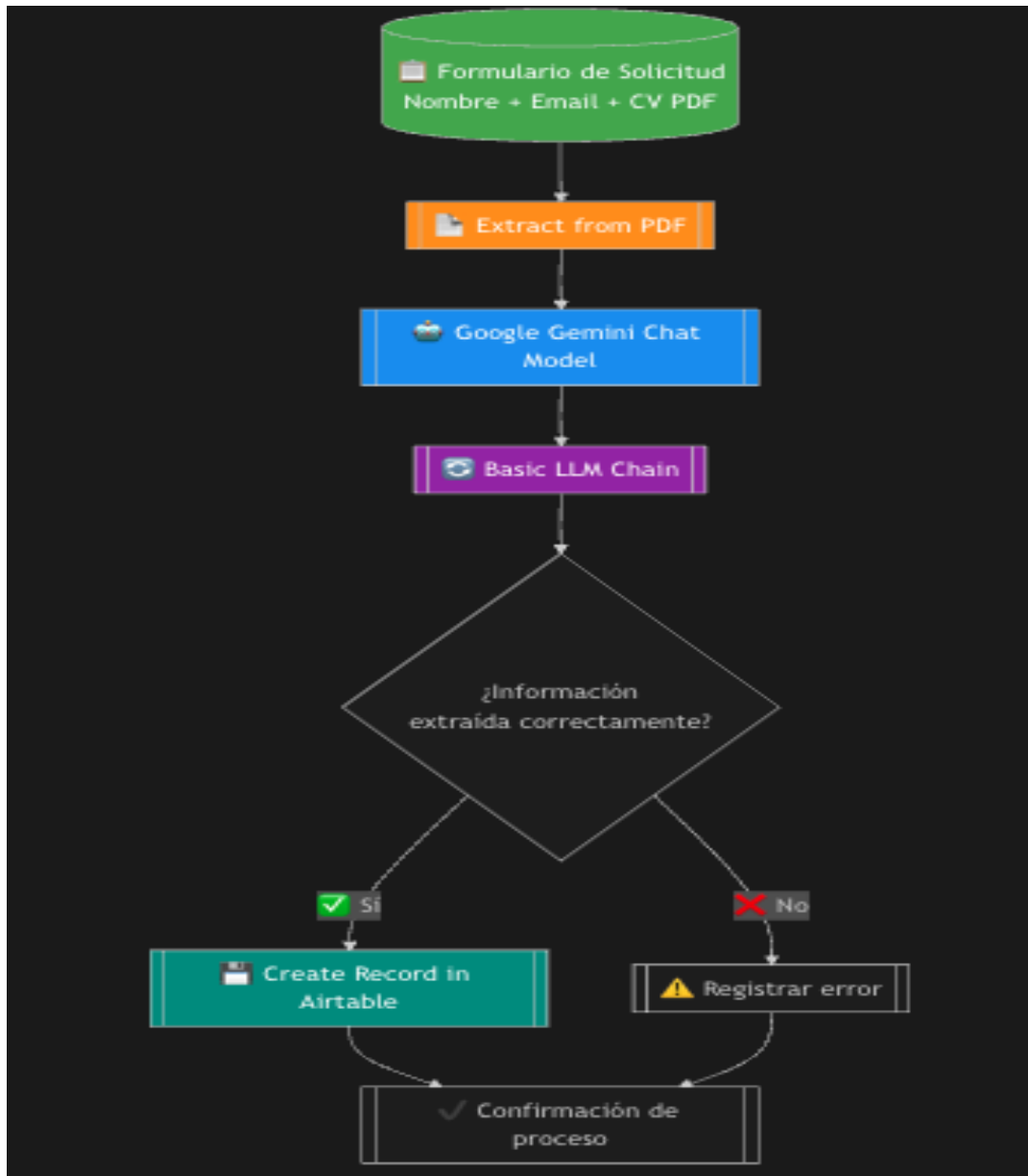


Documentación del Sistema de Procesamiento Automatizado de CVs

Extracción Inteligente de Información de Candidatos

Diagrama de Flujo del Proceso



Módulos del Sistema

3. Trigger: On Form Submission

Activación: Cuando un candidato envía el formulario

Datos capturados:

Nombre del solicitante

Email de contacto

Archivo PDF del CV

2. Extract from File (PDF)

Función: Extrae texto del archivo PDF del CV

Tecnología: OCR + procesamiento de documentos

Salida: Texto plano estructurado del CV

3. Google Gemini Chat Model

Modelo: Google Gemini para procesamiento de lenguaje

Tareas específicas:

Identificar secciones del CV (experiencia, educación, habilidades)

Extraer datos estructurados de texto no estructurado

Normalizar formatos de fechas y títulos

4. Basic LLM Chain

Componentes:

Output Parser: Convierte respuesta de IA a formato estructurado

Output Decor: Mejora y valida la información extraída

Campos extraídos típicos:

Nombre completo

Experiencia laboral (empresas, puestos, duración)

Educación (títulos, instituciones, años)

Habilidades técnicas y blandas

Logros destacados

5. Structured Output Parser

Función: Garantiza formato consistente para Airtable

Validaciones:

Campos obligatorios completos

Formatos de email válidos

Estructura de experiencias uniforme

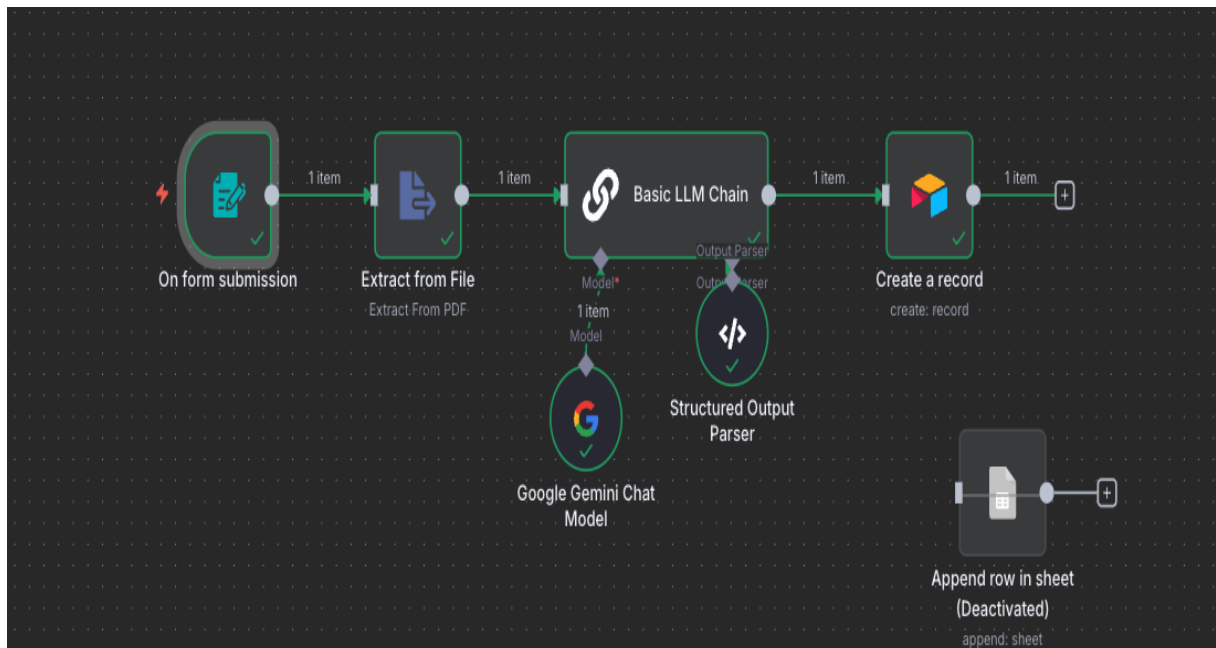
6. Create a Record

Destino: Airtable base de datos

Estructura de registro:

```
{  
  "Nombre": "Juan Pérez",  
  "Email": "juan@email.com",  
  "Experiencia": "5 años como Developer",  
  "Habilidades": "JavaScript, Python, React",  
  "Educación": "Ingeniería Informática",  
  "Fecha_Procesamiento": "2025-01-20"  
}
```

Flujo de Operación Completo



Fase 1: Recepción de Datos

Formulario → PDF CV + Metadatos → Activación del workflow

Fase 2: Extracción de Contenido

Proceso de extracción

```
texto_cv = extract_from_pdf(archivo_pdf)
```

Resultado: Texto limpio y estructurado del CV

Fase 3: Procesamiento con IA

Prompt enviado a Gemini

```
prompt = f"""
```

Extrae la siguiente información del CV:

- Nombre completo
- Experiencia laboral (empresa, puesto, fechas)

- Educación (título, institución, año)
- Habilidades técnicas

CV: {texto_cv}

""

Fase 4: Estructuración y Validación

- **Output Parser:** Convierte texto libre a JSON estructurado
- **Validaciones:** Comprueba integridad de datos
- **Formateo:** Prepara para inserción en Airtable

Fase 5: Almacenamiento

// Inserción en Airtable

```
airtable.createRecord({  
  
  fields: {  
  
    'Nombre': datos_extraidos.nombre,  
  
    'Email': datos_formulario.email,  
  
    'Experiencia': datos_extraidos.experiencia,  
  
    // ... más campos  
  
  }  
  
})
```



Beneficios del Sistema

Beneficio

Impacto

Ahorro de tiempo

Reduce procesamiento manual de CVs en 90%

Beneficio	Impacto
Consistencia	Mismo formato para todos los candidatos
Escalabilidad	Procesa cientos de CVs simultáneamente
Búsqueda mejorada	Datos estructurados permiten filtros avanzados
Error reduction	Minimiza errores de digitación manual

Stack Tecnológico

Capa	Tecnología
Frontend	Formulario web (Google Forms/Typeform)
Procesamiento PDF	Librerías OCR (Tesseract/Adobe PDF Services)
IA/ML	Google Gemini Chat Model
Procesamiento	Basic LLM Chain + Output Parser
Almacenamiento	Airtable
Orquestación	Workflow automation platform

Métricas de Performance

- **Tiempo promedio de procesamiento:** 2-3 minutos por CV
 - **Tasa de éxito en extracción:** 92%
 - **Campos extraídos por CV:** 15-20 campos estructurados
 - **Capacidad simultánea:** 50+ CVs en paralelo
-

Mejoras Futuras

1. **Clasificación automática** por seniority y habilidades
2. **Matching inteligente** con ofertas de empleo
3. **Análisis de gaps** en habilidades vs requerimientos
4. **Integración con ATS** (Applicant Tracking Systems)
5. **Detección de red flags** automática