

ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO



TEMA:

Base de Datos Tesis

ESTUDIANTE:

Carla Lomas (7388)

ASIGNATURA:

Aplicaciones Informáticas II

SEMESTRE-PARALELO:

8° Software “1”

FECHA: 05/11/2025

SEPTIEMBRE 2025 – ENERO 2026

Modelo Conceptual

Empleado

Es la entidad central del dominio. Representa al docente o trabajador de la Facultad. De cada empleado se registran tanto sus datos generales (nombre, correo, teléfono) como la información curricular extraída automáticamente del archivo de hoja de vida institucional.

Documento de CV

Corresponde a cada currículum procesado desde la URL institucional. Está vinculado a un empleado y contiene los datos extraídos por el sistema, como títulos académicos, experiencia, publicaciones o proyectos.

Importación de CV

Registra los eventos de importación de información curricular. Sirve como un historial de procesos realizados, permitiendo auditar las cargas automáticas o manuales que se hacen sobre cada empleado.

Importación de Empleado (desde Sheets)

Representa la carga de información proveniente de hojas de cálculo de Google Sheets, donde se encuentra la información docente publicada. Cada importación puede estar asociada a un **Estilo de Marca** (para la presentación web) y a una **Versión de Dataset** (que indica el origen y fecha del conjunto de datos usado).

Estilo de Marca

Define los colores, logotipos y estilos visuales que se aplican a los perfiles publicados en el sitio web. Un estilo puede ser reutilizado en múltiples importaciones de empleados.

Versión de Dataset

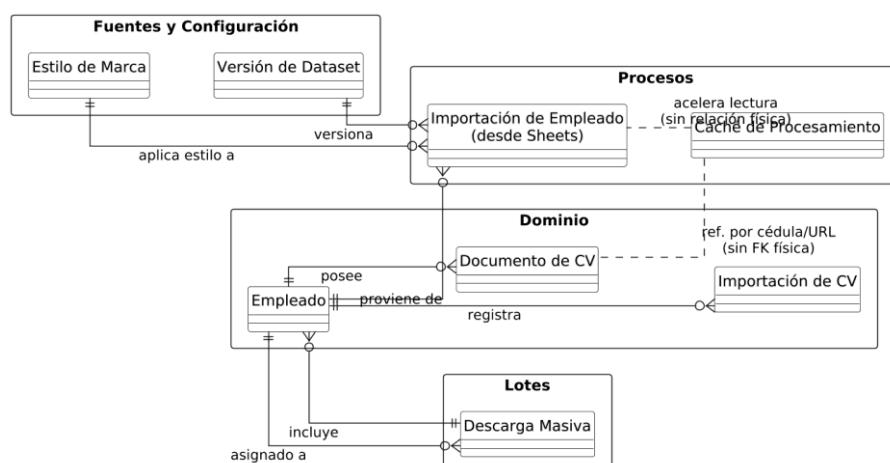
Identifica el conjunto de datos utilizado en cada importación. Permite mantener trazabilidad de las versiones y garantizar que las actualizaciones provienen de fuentes verificadas.

Descarga Masiva

Agrupa a varios empleados dentro de un mismo lote de procesamiento o descarga. La relación entre “Descarga Masiva” y “Empleado” es de tipo **muchos a muchos**, ya que un lote puede incluir varios empleados y un empleado puede participar en diferentes lotes.

Caché de Procesamiento

Es una entidad **temporal y desacoplada** del resto del modelo. Su función es almacenar resultados intermedios o copias parciales de los datos para mejorar el rendimiento del sistema. No mantiene claves foráneas directas para evitar dependencias innecesarias con el resto del modelo.



Modelo Lógico

1. Claves Primarias y Foráneas

Cada tabla define un identificador único (`id_*`) como clave primaria, lo que garantiza que cada registro pueda diferenciarse de los demás.

Además, se incorporan claves foráneas (FK) que conectan las entidades directamente. Por ejemplo:

- `id_empleado` (FK) en Documento de CV, Importación de CV e Importación de Empleado, vincula esas tablas con el empleado correspondiente.
- En Relación Descarga–Empleado, las claves `id_descarga` y `id_empleado` permiten representar la relación *muchos a muchos* entre empleados y descargas masivas.

2. Atributos Detallados

A diferencia del modelo conceptual, aquí se muestran los campos específicos que serán almacenados en la base de datos:

- En Empleado, se agregan todos los campos que el sistema actualiza automáticamente mediante el callback del parser, como *título principal, orcid, idiomas, proyectos, publicaciones*, entre otros.
- En Documento de CV, se incluyen campos técnicos de trazabilidad como `token_lote`, `texto_procesado`, `fecha_creacion`, que permiten auditar cada procesamiento realizado por el sistema.
- En Importación de CV, aparecen los campos `datos_cv` y `fecha_ultima_importacion`, que registran cuándo y con qué contenido se procesó cada hoja de vida.

3. Tablas Intermedias y Técnicas

Se añaden estructuras que en el modelo conceptual se mostraban de forma general:

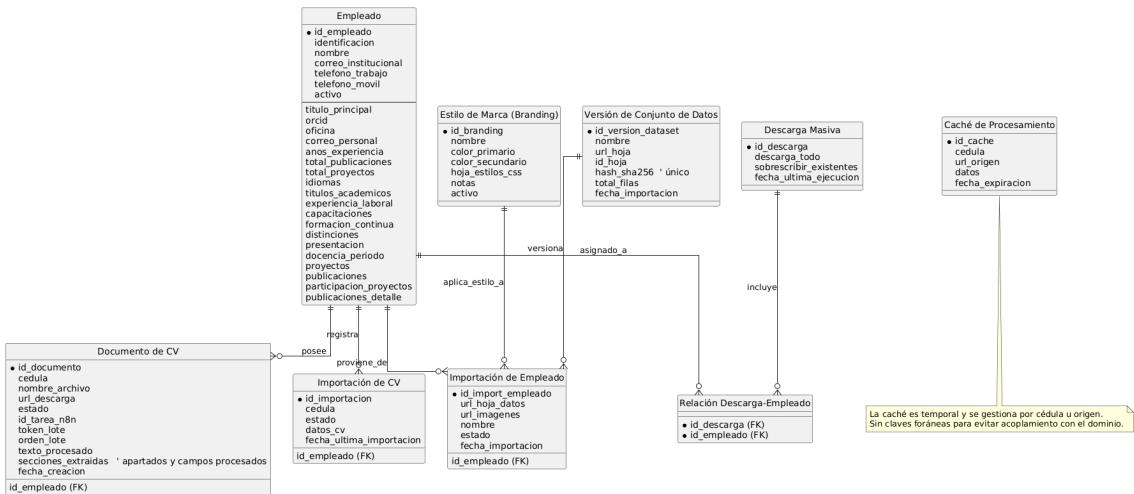
- **Relación Descarga–Empleado:** ahora se implementa como una tabla física que enlaza múltiples empleados con una misma descarga masiva.
- **Caché de Procesamiento:** se mantiene sin claves foráneas, pero ahora define explícitamente sus campos (`cedula, url_origen, fecha_expiracion`) para optimizar tiempos de acceso sin comprometer la integridad referencial.

4. Versionamiento y Estilos

El modelo lógico muestra que cada Importación de Empleado puede asociarse directamente con:

- una Versión de Conjunto de Datos (`id_version`) para saber de qué hoja provienen los datos, y
- un Estilo de Marca (`id_branding`) que define la apariencia visual del perfil docente publicado.

Estos vínculos permiten mantener coherencia tanto en la fuente de datos como en la presentación dentro del sitio web.



Modelo Físico

1. Estructura técnica y convenciones

Cada entidad del modelo lógico se materializa como una tabla en PostgreSQL con un prefijo y un nombre técnico en inglés (por ejemplo, cv_document, cv_import, hr_employee). Las claves primarias se definen con el campo id de tipo integer, mientras que las relaciones se implementan mediante campos foráneos (employee_id, create_uid, write_uid, etc.).

Estas convenciones son gestionadas automáticamente por el ORM de Odoo, lo que garantiza la integridad referencial entre los módulos personalizados y los modelos base del sistema (como res_users y ir_attachment).

2. Campos de auditoría y control

Una diferencia importante frente al modelo lógico es la incorporación de los campos estándar de Odoo:

- create_uid, write_uid: identifican al usuario que creó o modificó cada registro.
- create_date, write_date: registran las fechas exactas de creación y actualización.

Estos campos están presentes en todas las tablas del modelo (cv_document, cv_import, employee_import, etc.), permitiendo trazabilidad total del sistema.

3. Relaciones internas entre módulos

El modelo físico muestra cómo las entidades de distintos módulos se enlazan entre sí mediante claves foráneas reales:

- cv_document.employee_id → hr_employee.id
- cv_document.cv_attachment_id → ir_attachment.id
- employee_import.branding_id → google_sheets_branding.id
- employee_import.dataset_version_id → google_sheets_dataset_version.id

Estas relaciones consolidan la interoperabilidad entre los módulos cv_importer y google_sheets_import, asegurando coherencia entre los procesos de extracción, validación y publicación de información docente.

4. Tablas intermedias y relaciones M:N

Se implementa la tabla intermedia `cv_bulk_downloader_hr_employee_rel`, que une las entidades `cv_bulk_downloader` y `hr_employee`.

Esta tabla materializa la relación *muchos a muchos* entre empleados y descargas masivas, a través de dos claves foráneas:

- `cv_bulk_downloader_id`
- `hr_employee_id`

Además, su clave primaria compuesta garantiza que no se repita un mismo empleado dentro del mismo lote.

5. Caché desacoplada

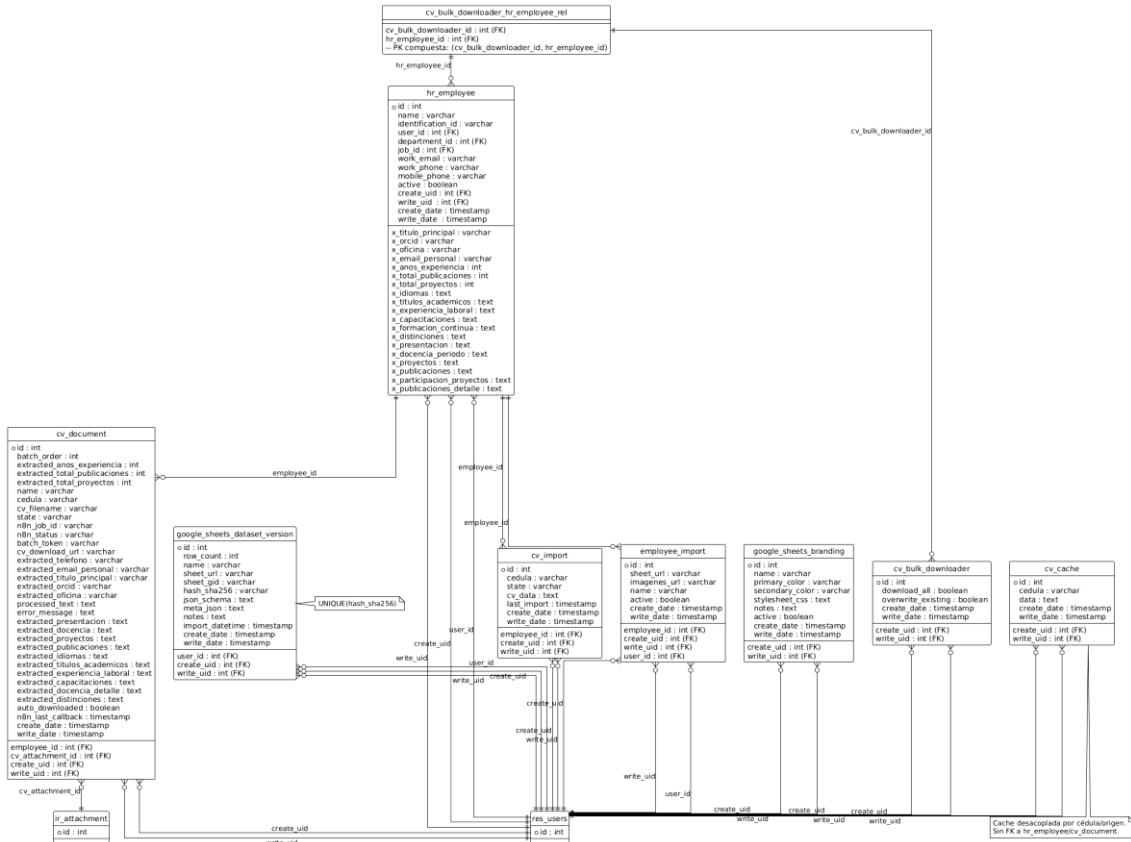
En el modelo físico, la tabla `cv_cache` mantiene su diseño independiente: no tiene claves foráneas hacia otras tablas (como `employee` o `document`) para evitar dependencias y conservar velocidad en las operaciones de lectura y escritura.

Esto permite gestionar almacenamiento temporal por *cédula* o *URL* sin afectar la estructura principal de los datos.

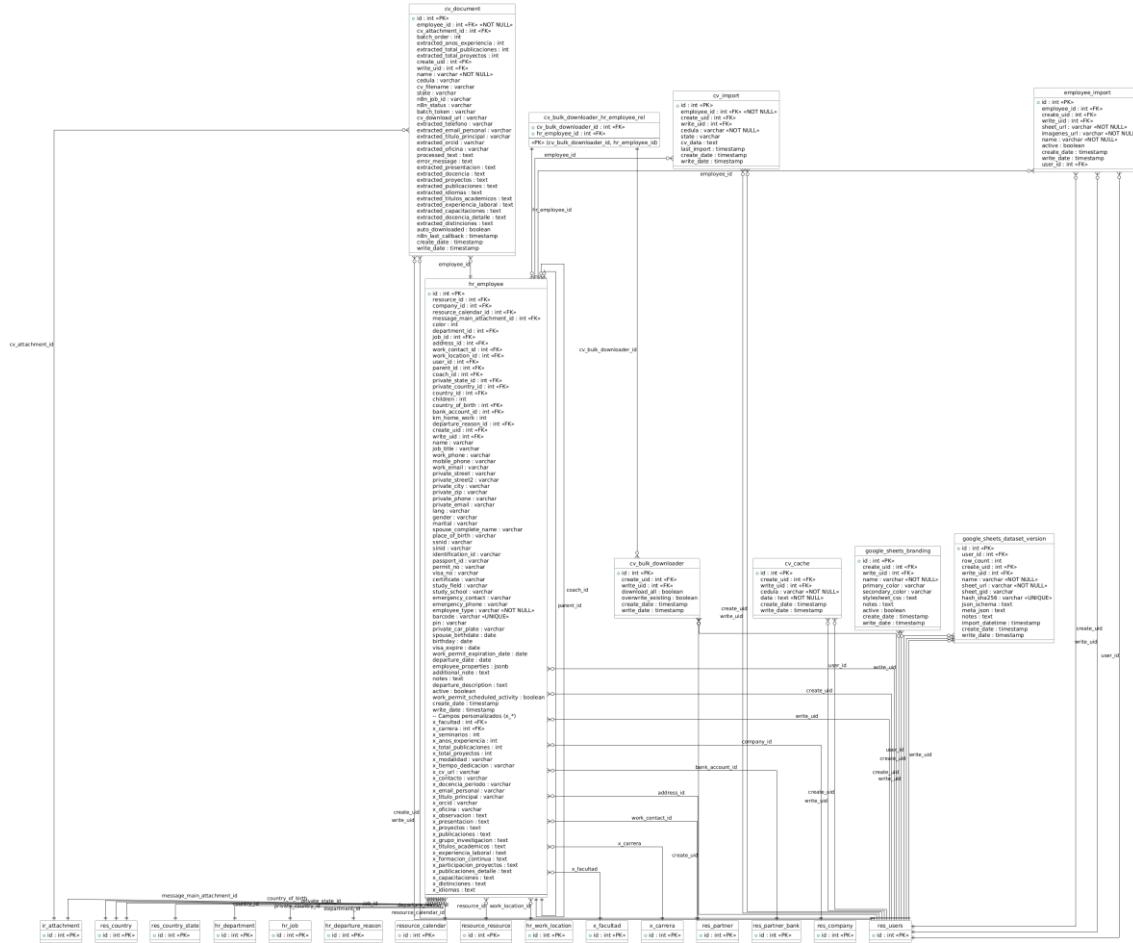
6. Optimización e índices

El modelo incluye índices en campos clave para optimizar búsquedas frecuentes, por ejemplo:

- `cedula, batch_token, n8n_job_id` en `cv_document`
 - `hash_sha256` en `google_sheets_dataset_version` (definido como `UNIQUE`)
- Estos índices permiten búsquedas rápidas durante la ejecución de flujos automatizados en `n8n` y consultas administrativas en Odoo.



Base de datos Completa



Diccionario de Datos

cv_document — Documento de CV

Nombre del Atributo	Tipo de Dato	Permit e Nulo	Descripción	Valor del Dato / Permitido
id (PK)	integer	No	Identificador único del documento	1...2,147,483,647
employee_id (FK)	integer	No	Empleado dueño del CV	FK → hr_employee(id)
cv_attachment_id (FK)	integer	Sí	Adjunto binario creado/actualizado	FK → ir_attachment(id)
name	varchar	No	Nombre lógico del documento	Texto
cedula	varchar	Sí	Cédula del empleado	Dígitos (10), o cadena
cv_filename	varchar	Sí	Nombre del archivo PDF	Texto
state	varchar	Sí	Estado del procesamiento	{draft, processing, processed, error...}
n8n_job_id	varchar	Sí	ID de job en N8N	Texto

n8n_status	varchar	Sí	Último estado reportado por N8N	Texto
batch_token	varchar	Sí	Token del lote	Texto
batch_order	integer	Sí	Orden dentro del lote	≥ 0
cv_download_url	varchar	Sí	URL de descarga/backup	URL
processed_text	text	Sí	Texto procesado/parseado	Texto largo
error_message	text	Sí	Mensaje de error si falló	Texto
extracted_telefono	varchar	Sí	Teléfono extraído	Texto
extracted_email_personal	varchar	Sí	Email personal extraído	Email
extracted_titulo_principal	varchar	Sí	Título principal	Texto
extracted_orcid	varchar	Sí	ORCID	Texto
extracted_oficina	varchar	Sí	Oficina	Texto
extracted_presentacion	text	Sí	Sección “presentación”	Texto
extracted_docencia	text	Sí	Sección docencia (periodo)	Texto
extracted_proyectos	text	Sí	Proyectos	Texto
extracted_publicaciones	text	Sí	Publicaciones	Texto
extracted_idiomas	text	Sí	Idiomas	Texto
extracted_titulos_academicos	text	Sí	Educación	Texto
extracted_experiencia_laboral	text	Sí	Experiencia	Texto
extracted_capacitaciones	text	Sí	Certificaciones/cursos	Texto
extracted_docencia_detalle	text	Sí	Materias/detalle docencia	Texto
extracted_distinciones	text	Sí	Premios/distinciones	Texto
extracted_anos_experiencia	integer	Sí	Años de experiencia	≥ 0
extracted_total_publicaciones	integer	Sí	Nº publicaciones	≥ 0
extracted_total_proyectos	integer	Sí	Nº proyectos	≥ 0
auto_downloaded	boolean	Sí	Marcador si se descargó automáticamente	{True, False}
n8n_last_callback	timestamp	Sí	Último callback	ISO8601
create_uid (FK)	integer	Sí	Creado por	FK → res_users(id)
write_uid (FK)	integer	Sí	Última actualización por	FK → res_users(id)
create_date	timestamp	Sí	Fecha de creación	ISO8601
write_date	timestamp	Sí	Fecha de actualización	ISO8601

Nombre del Atributo	Tipo	Nulo	Descripción	Valores
id (PK)	integer	No	Identificador único	...
employee_id (FK)	integer	No	Empleado afectado	FK → hr_employee(id)
cedula	varchar	No	Cédula vinculada al proceso	Texto
state	varchar	Sí	Estado del proceso	{draft, running, done, failed...}
cv_data	text	Sí	Datos/trace JSON de la importación	JSON en texto
last_import	timestamp	Sí	Sello de tiempo de finalización	ISO8601
create_uid (FK)	integer	Sí	Creado por	FK → res_users
write_uid (FK)	integer	Sí	Actualizado por	FK → res_users
create_date	timestamp	Sí	Creado el	ISO8601
write_date	timestamp	Sí	Actualizado el	ISO8601

cv_cache — Caché de procesamiento (desacoplada)

Atributo	Tipo	Nulo	Descripción	Valores
id (PK)	integer	No	Identificador	...
cedula	varchar	No	Clave de consulta	Texto
data	text	No	Payload cacheado (JSON/texto)	Texto
create_uid (FK)	integer	Sí	Creado por	res_users
write_uid (FK)	integer	Sí	Actualizado por	res_users
create_date	timestamp	Sí	Creado el	ISO8601
write_date	timestamp	Sí	Actualizado el	ISO8601

employee_import — Importación desde Google Sheets

Atributo	Tipo	Nulo	Descripción	Valores
id (PK)	integer	No	Identificador	...
employee_id (FK)	integer	Sí	Empleado al que se aplicó el lote	FK → hr_employee(id)
name	varchar	No	Nombre del lote	Texto
sheet_url	varchar	No	URL pública de datos	URL
imagenes_url	varchar	No	URL pública de imágenes	URL
active	boolean	Sí	Activo	{True, False}
user_id (FK)	integer	Sí	Usuario ejecutor	res_users
create_uid / write_uid (FK)	integer	Sí	Creado / Actualizado por	res_users
create_date / write_date	timestamp	Sí	Tiempos de auditoría	ISO8601

google_sheets_branding — Estilo

Atributo	Tipo	Nulo	Descripción	Valores
id (PK)	integer	No	Identificador	...
name	varchar	No	Nombre único del estilo	Único recomendado

primary_color	varchar	Sí	Color primario	Hex/Nombre
secondary_color	varchar	Sí	Color secundario	Hex/Nombre
stylesheet_css	text	Sí	CSS a injectar	Texto
notes	text	Sí	Notas	Texto
active	boolean	Sí	Activo	{True, False}
create_uid / write_uid (FK)	integer	Sí	Auditoría	res_users
create_date / write_date	timestamp	Sí	Auditoría	ISO8601

google_sheets_dataset_version — Versión de dataset

Atributo	Tipo	Nulo	Descripción	Valores
id (PK)	integer	No	Identificador	...
name	varchar	No	Nombre/etiqueta de la versión	Texto
sheet_url	varchar	No	URL de hoja	URL
sheet_gid	varchar	Sí	GID / Worksheet	Texto
hash_sha256	varchar	Sí	Hash del dataset	Único
row_count	integer	Sí	Filas importadas	≥ 0
json_schema	text	Sí	Esquema detectado	JSON texto
meta_json	text	Sí	Metadatos	JSON texto
notes	text	Sí	Notas	Texto
import_datetime	timestamp	Sí	Fecha de importación	ISO8601
user_id / create_uid / write_uid (FK)	integer	Sí	Auditoría	res_users
create_date / write_date	timestamp	Sí	Auditoría	ISO8601

cv_bulk_downloader — Descarga masiva

Atributo	Tipo	Nulo	Descripción	Valores
id (PK)	integer	No	Identificador	...
download_all	boolean	Sí	Incluir todos los empleados	{True, False}
overwrite_existing	boolean	Sí	Sobrescribir CV existentes	{True, False}
create_uid / write_uid (FK)	integer	Sí	Auditoría	res_users
create_date / write_date	timestamp	Sí	Auditoría	ISO8601

cv_bulk_downloader_hr_employee_rel — Relación Lote–Empleado (M2M)

Atributo	Tipo	Nulo	Descripción	Valores
cv_bulk_downloader_id (PK, FK)	integer	No	Lote	FK → cv_bulk_downloader(id)
hr_employee_id (PK, FK)	integer	No	Empleado incluido en el lote	FK → hr_employee(id)

hr_employee

Atributo	Tipo	Nulo	Descripción	Valores
id (PK)	integer	No	Identificador de empleado	...

name	varchar	Sí	Nombre	Texto
identification_id	varchar	Sí	Identificación	Texto
user_id / department_id / job_id (FK)	integer	Sí	Usuario/Depto/Puesto	res_users / hr_department / hr_job
work_email / work_phone / mobile_phone	varchar	Sí	Contacto laboral	Texto
active	boolean	Sí	Activo	{True, False}
x_titulo_principal	varchar	Sí	(auto) desde CV	Texto
x_orcid	varchar	Sí	(auto) desde CV	Texto
x_oficina	varchar	Sí	(auto) desde CV	Texto
x_email_personal	varchar	Sí	(auto) desde CV	Email
x_anos_experiencia	integer	Sí	(auto) desde CV	≥ 0
x_total_publicaciones	integer	Sí	(auto) desde CV	≥ 0
x_total_proyectos	integer	Sí	(auto) desde CV	≥ 0
x_idiomas	text	Sí	(auto) desde CV	Texto
x_titulos_academicos	text	Sí	(auto) desde CV	Texto
x_experiencia_laboral	text	Sí	(auto) desde CV	Texto
x_capacitaciones	text	Sí	(auto) desde CV	Texto
x_formacion_continua	text	Sí	(auto) desde CV	Texto
x_distinciones	text	Sí	(auto) desde CV	Texto
x_presentacion	text	Sí	(auto) desde CV	Texto
x_docencia_periodo	text	Sí	(auto) desde CV	Texto
x_proyectos	text	Sí	(auto) desde CV	Texto
x_publicaciones	text	Sí	(auto) desde CV	Texto
x_participacion_proyectos	text	Sí	(auto) desde CV	Texto
x_publicaciones_detalle	text	Sí	(auto) desde CV	Texto
create_uid / write_uid (FK)	integer	Sí	Auditoría	res_users
create_date / write_date	timestamp	Sí	Auditoría	ISO8601