

# Collection Report (Notas y solución a los problemas)

- **Interpretación de campos ambiguos**
  - Se documentaron múltiples variables con nombres poco intuitivos (e.g., `Última OCC (x 1000)`, `Éxito IO (h)`).
  - Se realizaron consultas directas con el equipo docente y revisión cruzada entre archivos para inferir su significado.
  - Se creó un [Diccionario de datos](#) en Notion para uso interno del equipo, sujeto a validación por expertos.
- **Estrategia de validación de calidad**
  - Se diseñaron varios script en Jupyter -Python para evaluar métricas básicas de calidad: completitud, datos usables y relaciones entre datos.
  - Se identificaron duplicados y valores nulos en campos clave.
- **Control de versiones y trazabilidad**
  - Se implementó un sistema de control de versiones para registrar cada modificación de los archivos originales.
  - Se documentaron los pasos de transformación en el archivo de Jupyter de manera técnica para reproducibilidad.