


# Digital Geology: Cross-Checking Legacy Well Data

Legacy well files read by AI agents, every conflict flagged, every call made by a geologist.

**Status.** A completed pilot on a producing-field project in the Volga-Ural basin, December 2025 – February 2026. It ran within the project's data rules: the models ran locally, and any external model saw only anonymised data.

The question

## Decades of well data sit in scans and paper files: wellhead coordinates, deviation surveys, picks and core descriptions. Documents for the same well often disagree, and a model built on them inherits the conflicts.

The question is which values can be trusted before they go into the model, and how long it takes to find out.

This note is for teams preparing a geological model, a field review or a new well plan on fields with a long history.

Approach

A small team of AI agents reads well files of different ages and formats, compares the values for each well across documents and flags every conflict. Each conflict goes to a geologist, who decides which document is right.

Every value stays traceable to its source document. The agents do the reading; the call and the signature stay with the geologist.

### Read

Wellhead coordinates, deviation surveys, picks and core descriptions taken from scans and well files

### Cross-check

Values for the same well compared across documents; every conflict flagged

### Decide

The geologist resolves each conflict; the result is one well database with a source for every value

What this is based on

Pilot on a producing-field project in the Volga-Ural basin, December 2025 – February 2026: legacy well records, including wellhead coordinates, deviation surveys, picks and core descriptions.

Result: legacy-data preparation went from months to weeks (expert-validated), with every value traceable to its source document.

These figures describe one pilot. Results on another field depend on its archive.

| Stages and outputs | Work | Output |
|---|---|---|
| **1. Archive review** | Document types, scan quality and target fields reviewed. Data rules agreed with the operator: where the models run and what is anonymised. | Archive inventory and agreed data rules |
| **2. Reading** | Agents extract values from each document, with a link to the source page. | Values with their sources |
| **3. Cross-check** | Values compared between documents; conflicts flagged. | Conflict register |
| **4. Geologist review** | Each conflict resolved by the geologist, with the reason recorded. | One well database, checked by a geologist |
| **5. Hand-over** | Database delivered in the project's format, ready for the model and well planning. | Database and a short report |

Data the work relies on

The fields or wells in scope and the decision the data is for.

Scans and digital files of well records: coordinates, deviation surveys, tops and core descriptions.

The project's data rules: where data may be processed and what must be anonymised.

The target format of the database.

One technical contact who knows the archive.

## For discussion

This note is written to be printed and discussed by the team. A first conversation is about the archive and what the data is needed for; it needs no confidential data.

Data stay under the operator's rules: the models run locally, and external models see only anonymised data.

Client and employer data are never shown.

**maxim@balaguroff.com**

Maxim Balagurov, senior petroleum geologist with 15+ years in petroleum geoscience (since 2009). Member of EAGE and SPE.