

98% field accuracy.

Is the document right?

Measure whole-document success and the fields that can change the decision.



THE OPERATING QUESTION

98%

correct fields

Illustrative: 980 of 1,000 fields

01 / CHANGE THE UNIT OF EVALUATION

200 fictional documents; five fields each

98%

980 correct fields / 1,000 field decisions

90%

180 fully correct documents / 200 documents

20 field errors occur in 20 separate documents. One wrong field is enough to fail each affected document.

02 / BREAK OUT THE DOCUMENT GROUPS

Group	Correct / total	Rate
English	118 / 120	98.3%
Arabic	48 / 60	80.0%
Mixed	14 / 20	70.0%

Evaluate critical fields and complete documents.

Group scores need their own denominator and error review.

All evaluation figures are fictional. Each affected document has one field error. Language-group rates are an explanation of denominators, not measured model performance.