

Station 1 — fix a broken scan (OCR)

Source file: przyklad-korpus/poznan-1956/wycinki-prasowe/wycinek-01-ocr.txt

This is a press release from 1956 “run through bad OCR”: 1 instead of l, 0 instead of o, lost Polish characters, broken words.

Prompt (paste it, then paste the text from the file underneath)

Fix the OCR errors in the Polish text below. Restore correct Polish spelling with its diacritics (ą, ę, ł, ó, ż, ź...), fix the swapped characters (1→l, 0→o), and rejoin broken words. Keep the original meaning and style. Add nothing of your own, and don't comment.

What to expect: clean, readable release text in a few seconds — ready for further work.

Harder variant

After fixing the text, add a short analysis: what kind of language is this release written in (neutral / propaganda)? Point to 3 phrasings that show it.

Why it matters in research: old PDFs, archive scans, and photos of documents almost always have OCR errors. The model cleans them faster than manual correction — and you stay on verification.