
Lessons from your pdf2md

Placeholder

This deck arrives with your submissions

It is built from the lesson slides in your pdf2md writeups — collected after the deadline, triaged, and presented back to the class.

Until then it is deliberately empty: the content is the class's own work, and that does not exist yet.

Choose the table-recovery path

10 min

- Two designs tie on average score. A uses native PDF text first: fast on digital pages, but blind on raster-only pages.
- B OCRs every requested page: it handles scans, but is slower and sometimes splits spanning headers. Both retain page-and-box links.
- Choose the primary path and the exact fallback trigger. Name one challenge case that could reverse your design; compare with a pair.
- Deliverable: a four-line routing rule plus its challenge case.