

PDFHaul Challenge Set - Section A

Cross-Generator Equivalence Case

Document ID: PDFHAUL-CHALLENGE-A

This page contains identical wording rendered by three different PDF generation pipelines. A detector that compares raw file bytes or PDF object structure will see these as different files. A detector that compares rendered content should flag all copies as duplicates of each other.

PDFHaul Challenge Set - Section A

Cross-Generator Equivalence Case

Document ID: PDFHAUL-CHALLENGE-A

This page contains identical wording rendered by three different PDF generation pipelines. A detector that compares raw file bytes or PDF object structure will see these as different files. A detector that compares rendered content should flag all copies as duplicates of each other.

PDFHaul Challenge Set - Section A

Cross-Generator Equivalence Case

Document ID: PDFHAUL-CHALLENGE-A

This page contains identical wording rendered by three different PDF generation pipelines. A detector that compares raw file bytes or PDF object structure will see these as different files. A detector that compares rendered content should flag all copies as duplicates of each other.

PDFHaul Challenge Set - Section A

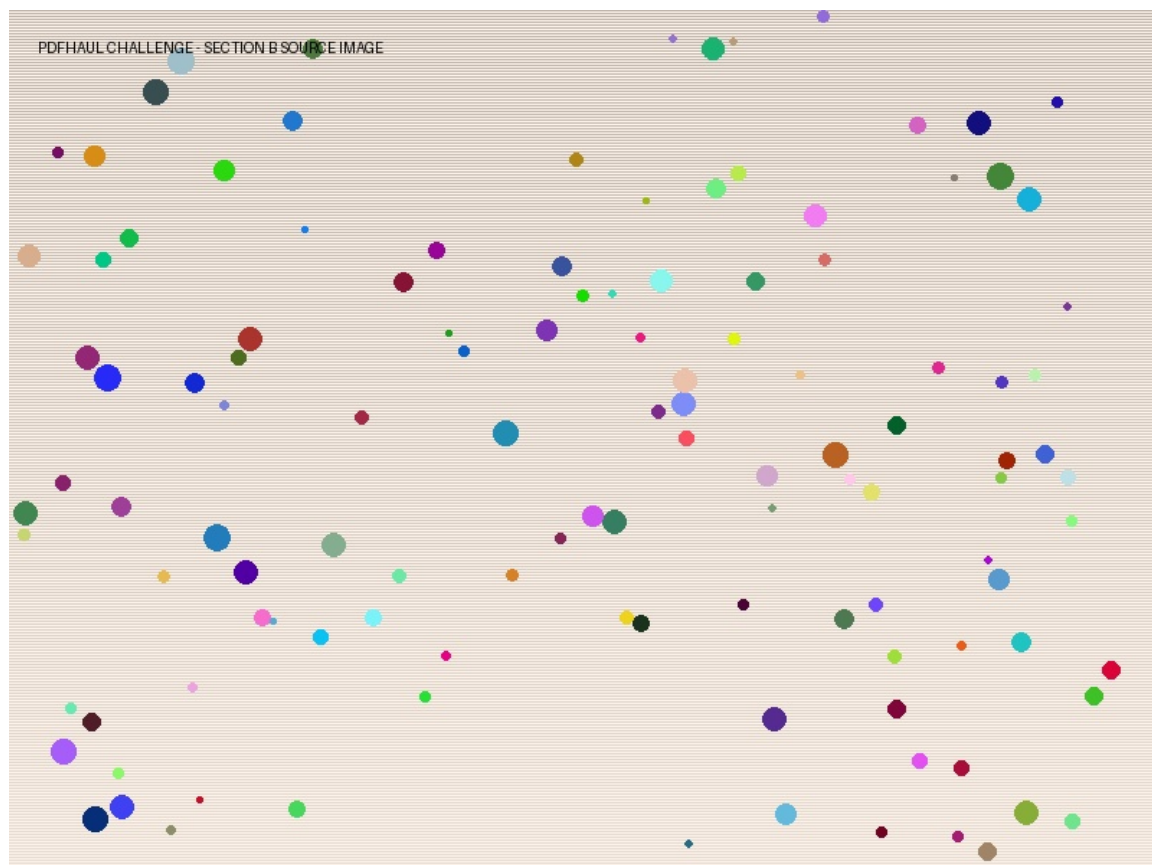
Cross-Generator Equivalence Case

Document ID: PDFHAUL-CHALLENGE-A-NONMATCH

This page is rendered by one of the same pipelines used above, but the wording is different. It must not be flagged as a duplicate of any Section A match page.

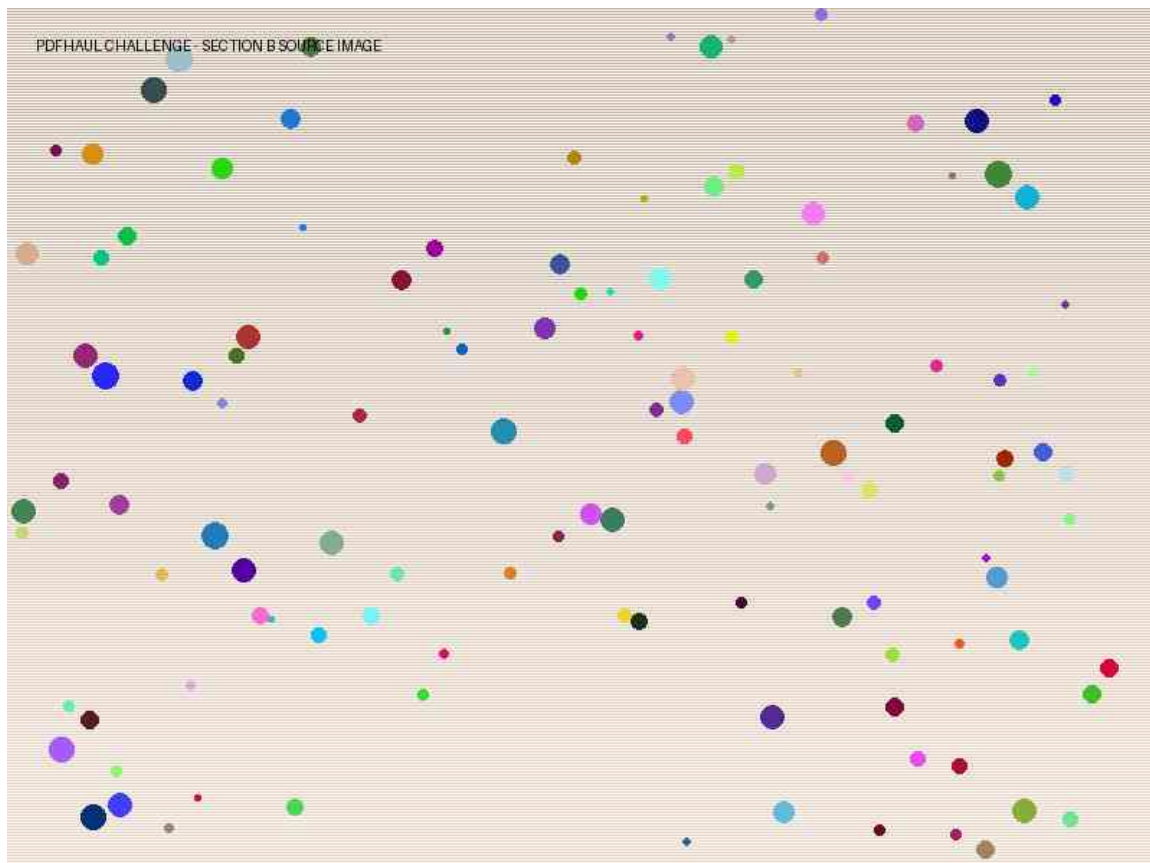
PDFHaul Challenge Set - Section B

Recompressed Image Equivalence Case



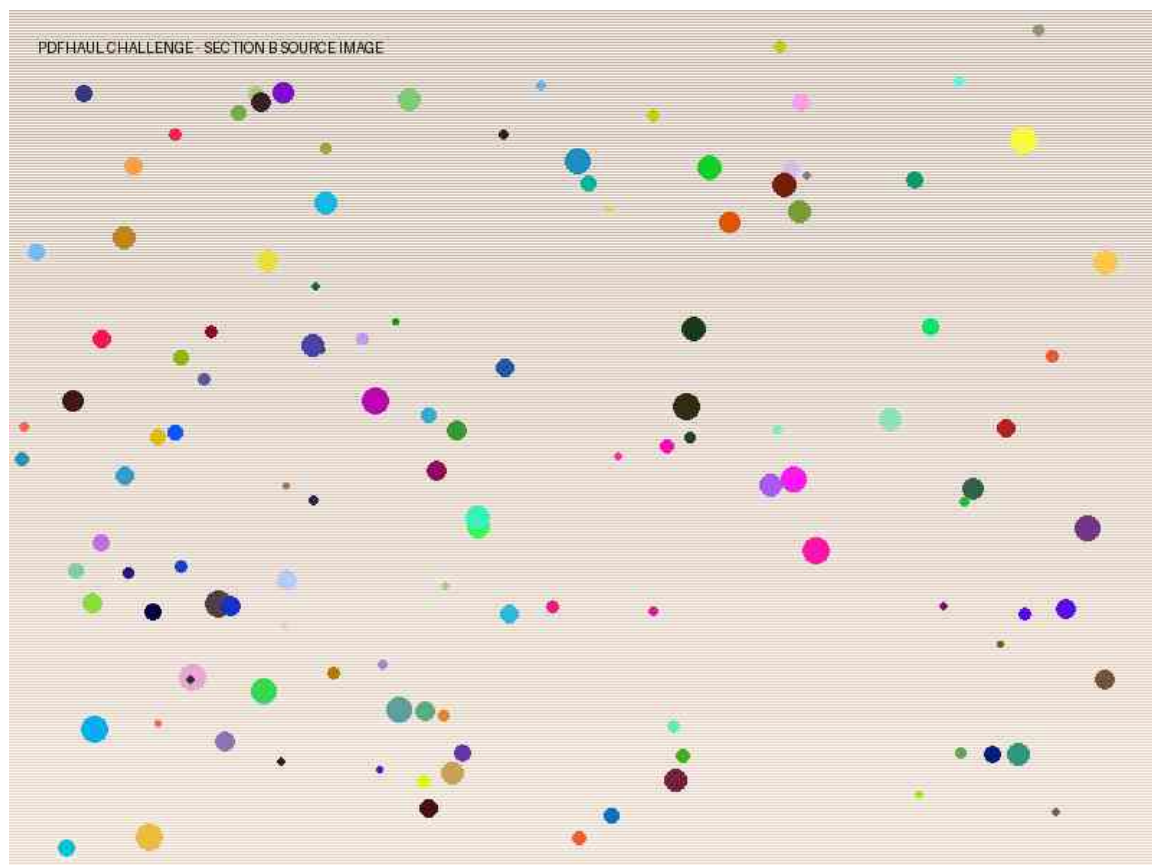
PDFHaul Challenge Set - Section B

Recompressed Image Equivalence Case



PDFHaul Challenge Set - Section B

Recompressed Image Equivalence Case



PDFHaul Challenge Set - Section C

Shift/Rotation Equivalence Case

Quarterly Compliance Summary - Section C test content for shift and rotation detection.

PDFHaul Challenge Set - Section C

Shift/Rotation Equivalence Case

Quarterly Compliance Summary - Section C test content for shift and rotation detection.

PDFHaul Challenge Set - Section C

Shift/Rotation Equivalence Case

Quarterly Compliance Summary - Section C test content for shift and rotation detection.

PDFHaul Challenge Set - Section C

Shift/Rotation Equivalence Case

Annual Risk Assessment Overview - different content, same rotation as the Section C rotated match page.

PDFHaul Challenge Set - Section E

Visually Similar, Substantively Different

INVOICE

Invoice No: INV-1042

Bill To: Acme Consulting LLC

Service: Monthly PDF workflow subscription

Amount Due: \$100.00

Terms: Net 30

PDFHaul Challenge Set - Section E

Visually Similar, Substantively Different

INVOICE

Invoice No: INV-1043

Bill To: Acme Consulting LLC

Service: Monthly PDF workflow subscription

Amount Due: \$100.50

Terms: Net 30

PDFHaul Challenge Set - Section F

Simulated Rescan Equivalence Case

Employee Handbook - Section 4: Expense Reimbursement Policy

PDFHaul Challenge Set - Section F

Simulated Rescan Equivalence Case

Employee Handbook - Section 4: Expense Reimbursement Policy

PDFHaul Challenge Set - Section F

Simulated Rescan Equivalence Case

Employee Handbook - Section 9: Remote Work Guidelines