



# Remove duplicate media: unified workflow

**Hey! A quick word about Mediapapa first:** duplicate hunting by hand is painful, especially across images, PDFs and videos at once. Mediapapa detects exact duplicates by content hash and lets you replace files at the same URL, so no internal link breaks during cleanup.

## The five-step loop

---

Apply in order, for every cluster of duplicates.

- Detect:** hash, name/size/date triangulation, or repeated URL references.
- Classify:** group by format (image, PDF, video) and by role.
- Verify usage:** a duplicate can still be the one someone references.
- Consolidate:** choose a canonical file and update all references first.
- Delete safely:** trash, test downloads and playback, then purge.

## Images — extra rules

---

Never delete generated thumbnails directly. WordPress manages them.

- Confirm the file is a true duplicate, not a generated thumbnail (image-150x150.jpg).
- Keep the best-optimised original; let WordPress remove its thumbnails automatically.
- Update content references to point at the surviving URL before deleting.
- WooCommerce product images and variations verified separately.

## PDFs — version chaos

---

One canonical link per concept. Replace, do not duplicate.

- Canonical version chosen for each PDF concept (brochure, spec sheet, catalogue).
- Internal page links updated to point at the canonical URL.
- Email templates checked: marketing platform, confirmation emails, automated notifications.
- Email signatures audited (yours and your team's).
- External backlinks checked: keep the old file or set up a redirect if traffic is at stake.
- Old version deleted only after every reference is verified.

## Videos — hosting strategy first

---

Decide whether the video should be in WordPress at all.

- Decide between local upload and embed (YouTube, Vimeo) before deduplicating.
- If local: one resolution kept, multiples removed.
- If embed: one canonical URL used across every post that references the video.
- Caption and transcript files updated to match the surviving video.
- Playback tested on every page that references a deleted copy.

## After deletion

---

Quick sweep, before customers find the issue first.

- Every page that referenced removed files opened and scrolled.
- Download buttons clicked, files actually open.
- Test email sent for every automated email containing a media link.
- Video players load and play on each affected page.
- Mobile retested where breakages tend to surface first.
- Page cache cleared. CDN purged. Third-party file services purged.

### Hash-based duplicate detection, Safe Replace included.

Mediapapa Pro finds exact duplicates across images, PDFs and videos by content hash, then lets you consolidate references with Safe Replace: the canonical file keeps the URL, every page keeps working. No manual link replacement.

[See Mediapapa Pro plans →](#)

support@wp-mediapapa.com · wp-mediapapa.com