What is File Carving?

🧩 What is File Carving?

File Carving is a digital forensics technique used to **recover files from unstructured or damaged data sources**, without using file system metadata such as filenames, timestamps, or directory structure.

It is typically used when:

- The file system (e.g., FAT, NTFS, EXT) is corrupted or missing.
- Files have been **deleted intentionally**.
- A disk image or memory dump is being analyzed for **hidden or lost files**.

File carving is **content-aware**, meaning it relies on the actual **structure of file formats** (like headers and footers) to extract data.

🛠️ How File Carving Works – Step-by-Step

🔹 1. Raw Data Scanning

The tool scans the entire raw data of the storage medium (like a hard drive or disk image) **byte by byte**.

🔹 2. Header Identification

Each file format begins with a specific **signature** or **magic number**.
For example:

- JPEG → Starts with FFD8 (hex) → ÿØ
- PDF → Starts with %PDF
- ZIP → Starts with 504B0304 (hex)

🔹 3. Footer Identification

Some formats have footers to indicate the end of the file.
For example:

- JPEG → Ends with FFD9

Once a header is found, the tool keeps reading until:

- A matching footer is found, or
- A specified maximum file size or pattern is reached.

🔹 4. File Reconstruction

The tool extracts all bytes from header to footer (or predefined size) and saves them as a new file.

📂 Types of File Carving

✅ Header–Footer Carving

- Looks for known **start and end** patterns (e.g., JPEG: FFD8 to FFD9).
- Most accurate when file is not fragmented.

✅ Header–Max File Size Carving

- Used when file format has **no footer** (e.g., certain DOC files).
- Extracts a chunk of data from header up to a maximum limit.

✅ Content-Based or Semantic Carving

- Uses internal structure or patterns (e.g., EXIF data in images).
- Advanced but computationally heavy.

✅ Fragment Carving

- Tries to reassemble **fragmented files**.
- Requires deep knowledge of file structure (like ZIP central directory or MP4 frame headers).

🧪 Real-Life Example: Digital Forensics Case

🕵️ Scenario:

A suspect deletes illegal images before a device is seized.

👣 Investigation:

- Forensic expert creates a **disk image** (using tools like dd or FTK Imager).
- Uses PhotoRec to **carve files from raw sectors**.
- Tool scans for JPEG headers (FFD8) and footers (FFD9).

- Successfully **recovers deleted images**, even though file names and paths are gone.

These images can be used as **digital evidence in court**.

🧰 Tools Used for File Carving

Tool	Description
Scalpel	Fast and customizable; supports multiple file formats.
Foremost	Command-line based, originally developed by the U.S. Air Force.
PhotoRec	Powerful recovery tool for over 480 file formats; part of TestDisk suite.
Autopsy/Sleuth Kit	GUI-based forensic tool; includes file carving support.
Bulk Extractor	Extracts useful artifacts like credit card numbers, email addresses, etc., from raw disk images.

📌 Key Applications

- **Recovering deleted or lost files** from corrupted drives.
- **Digital forensics investigations** (e.g., in cybercrime, fraud).
- **Data breach analysis** (detect if sensitive files were exfiltrated).
- **Disaster recovery** when backup is not available.

❌ Limitations of File Carving

Limitation	Description
🔻 **Fragmentation Issues**	Carving tools often fail if a file is broken into non-contiguous parts.
🔻 **No File Metadata**	File name, creation date, original path are usually **not recoverable**.
🔻 **False Positives**	Tools may extract junk data that appears to be a file.
🔻 **Limited Format Support**	Only files with known headers/footers or internal structures can be carved.

🔐 File Signatures (Examples)

File Type	Header (Hex)	Footer (Hex)
JPEG	FFD8FFE0 or FFD8FFE1	FFD9
PNG	89504E47	49454E44AE426082
PDF	%PDF-	%%EOF
ZIP	504B0304	(no fixed footer)