

# Digitizing Information Worksheet

DRAFT VERSION FOR FEEDBACK

## Mechanism

Analog work is often useful at the point of creation: sketches and notes on paper, ideas jotted in notebooks, developing an idea on a whiteboard, using sticky notes, and so on. Even paper books have advantages — no Digital Rights Management, better formatting, no limits on access, and cheap second-hand availability.

Digital availability has different advantages: retrieval, reuse, backup, and easier integration into projects. Referring to information from a digital Idea Collection, or including the content directly, is much easier. You can place sketches on a page, link to relevant books, or pull a passage into a project note without hunting through piles of paper.

The intervention is to reduce the friction between analog creation and digital availability without creating an archive-maintenance hobby. This includes digitizing paper material, but also turning audio recordings into transcripts.

## Applicability

This worksheet is relevant if

- you work with analog materials — paper, whiteboards, sticky notes, books, etc. —

and want the information available digitally,

- you can actually work with digital material — especially with books, some people still prefer to read on paper.

Be careful if you spend more time digitizing information and managing an archive than actually using it. The drive to make it «complete» or «properly formatted» can easily become a time sink. The goal is to have ideas and other material available for creative work, not to become an archivist of your life.

## Intervention Variables

To ensure that you actually have digital information and can work with it, the following workflow requirements are relevant:

- **File Format:** Use widely supported, reasonably future-proof formats, e.g., PDF, markdown, JPG, PNG, or TIFF.
- **Backups:** Be careful with storage media — they all die, the question is when. See Worksheet Backups.
- **Satisficing quality:** File size usually increases with quality, meaning backups take longer and require larger media. Especially with ideas — sketches, notes, rough material — the goal is to work with the information. It is not an art repository. In practice, downsized images are often enough, e.g., reducing a high-quality smartphone photo to 1200 px. Check whether the quality is sufficient: work with a copy, downsize it, close the file, reopen it, and see whether it still works.
- **Keep the infrastructure ready:** The cap-

ture device, folder, app, or scanner must be ready when the material appears. If setup takes more than two minutes, digitization will likely remain occasional rather than integrated.

- **Manual transcription remains valid:** For short or idiosyncratic handwritten notes, typing by hand is often faster and more reliable than correcting OCR. Being able to type properly with ten fingers takes time to learn, especially if you developed your own typing style, but the speed gains are worth it — not only for digitizing information, but also for writing in general: emails, books, notes, and everything else.
- **Avoid backlogs:** Every digitization workflow needs a backlog limit. If more than 20 unprocessed scans accumulate, stop digitizing non-urgent material and process, discard, or file what is already captured.

There are many ways to digitize information, depending on the original material. Use only the module relevant to the material you are trialing.

## Audio Information

This includes audio notes, e.g., from a voice-recording app such as «Just Press Record», but also transcribing podcasts or videos to make the information available.

- **Audio Transcription Software:** Speech-to-text software turns audio recordings into text files. Whisper Transcription is a good option because it can run locally. Cloud transcription may create privacy

## Box 1: Rights Boundary

**Working with legally accessible digital text:** If a digital source allows copying or export, use the built-in export, citation, accessibility, or library functions first. For legally permitted short excerpts, screenshots can be sufficient. Do not use this workflow to bypass DRM, license terms, platform restrictions, or redistribution limits.

**Legal Aspects:** Only digitize material you are legally allowed to copy and store. Check local law, license terms, library rules, and platform terms. Personal backup, accessibility, excerpts, and research use can be treated differently depending on jurisdiction. Do not redistribute scanned or extracted material.

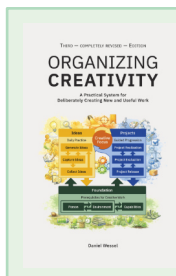
**Privacy/Sensitivity:** Decide whether cloud tools may process the material. For private, client, legal, medical, unpublished, or emotionally sensitive material, prefer local tools or manual transcription.

issues, depending on the material and service (see Box 1: Rights Boundary). Transcription quality depends on the model, the recording quality, and how clearly the speaker can be understood. Learning to speak clearly also has benefits outside of capturing ideas.

- **Manual Transcription:** Hardware and apps can assist manual transcription, e.g., a foot pedal to switch between play/pause and rewind a few seconds. Often, a simple audio player, such as QuickTime, plus a

## Relevant Chapters

For background information, see Chapter 5: Environment.



# Digitizing Information Worksheet

text file is enough. Switch between text and player with CMD + TAB on Mac or CTRL + TAB on Windows, and use SPACE to pause the audio application. This lets you transcribe without taking your hands off the keyboard. If you have many audio notes, transcription software saves a lot of time and energy.

## Notes on Paper

This includes ideas on notepads, in journals, and similar sources. Photos can be used, but searchable text is usually more useful. Images should be reserved for sketches and similar material.

- **Manual Transcription:** Especially with idiosyncratic handwriting, typing notes by hand is often easiest. It also allows you to add context where needed. Moving from capturing to collecting means the idea has to be understandable in the future, not only in the moment. Typing is often easier if the note is photographed first and transferred to the computer. Then you have the note and your text file directly in front of you, and the image can remain part of the backup until it is transcribed. The risk is backlog accumulation — if you transfer it to the computer, transcribe it soon.
- **Use Audio Transcription Software:** Reading a handwritten note aloud and using transcription software can save typing. However, you need the privacy to do so and must tolerate hearing your own voice read even bad ideas. Listening to music via headphones might help. Using audio notes in the first place might be

easier — see Worksheet Capturing Ideas.

- **Optical Character Recognition (OCR) supported by AI:** OCR (see Box 2) works best with typed text, e.g., from books, but can also work with handwriting. AI has made this more feasible, and some note-taking software includes it. However, AI can hallucinate or misread notes. For notes that will guide decisions, verify the output against the original before deleting anything.
- **Using a Digital Photo:** For sketches and similar material, simply photographing the note is often enough. Smartphone photos usually provide sufficient quality (see Box 3). In a pinch, even a webcam photo can be enough, e.g., using Photo Booth on Mac or taking a screen capture. Adding the current date to the photo is useful if the creation date later changes (see Figure 1 for a way to do that with macOS Automator, saved as a Service it is then available from the context menu as Quick Action).
- **Scanners:** If you have many paper notes on separate pages, a document scanner produces very good results. These scanners have a feed similar to a copier. If the material is not on separate pages, a flatbed scanner or book scanner (see Digitizing Books) is an option. You place the paper on a glass surface, which allows you to scan books and thicker material, but these scanners are usually slow.


## Digitizing Books

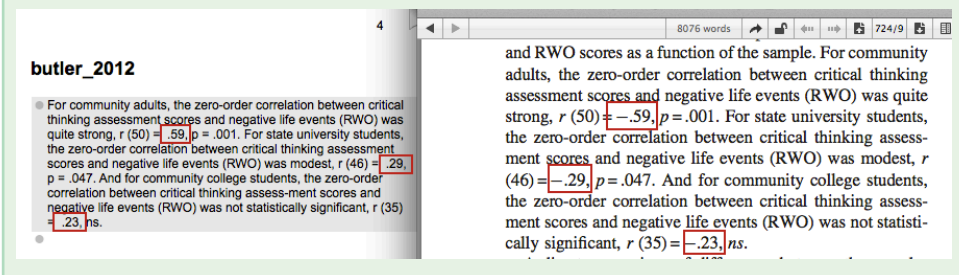
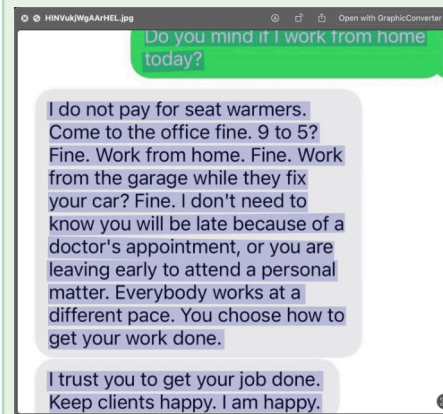
With books, the temptation is strong to go for a high-quality capture. But this is not about

preserving beauty; it is about functionality. If beauty is the main concern, nothing beats high-quality binding and print.

There are also legal issues that determine

### Box 2: Optical Character Recognition (OCR)

Optical Character Recognition, or OCR, turns text on an image into selectable and copyable text. Some operating systems do this automatically, e.g., current versions of macOS (see image below, the small  icon appears after a few seconds in Quick View, then the text in the image can be selected), and some PDF apps can do it as well.



## DRAFT VERSION FOR FEEDBACK

whether you may copy books, e.g., for personal backups, accessibility needs, library scans, excerpts, and similar cases. See Box 1: Rights Boundary.

However, bad scans or badly aligned text can produce mistakes, and apps differ in OCR quality. Try what works best for you and use good OCR; otherwise it is a pain in the ass to correct later (see the image below).

One cognitive process in particular, self-efficacy, is hypothesized to mediate the influence of intrapersonal factors on learning and subsequent behavior change. To increase self-efficacy, strategies to improve goal achievement are embedded throughout the intervention. These strategies include self-monitoring, intention formation, goal setting, goal review, feedback on performance, self-efficacy benefits, ball™, problem-solving, social support, and tailoring. Expanding on SCT, control theory (251) proposes that setting goals, monitoring behavior, receiving feedback, and reviewing relevant goals after obtaining feedback are central to self-management and behavioral control. Ecological theory (261) proposes a more comprehensive model of health behavior by suggesting that behaviors are embedded in four nested

If a document has bad OCR, consider running a better one over it. If that is not possible, exporting the pages as images, creating a new document from those images, and running OCR again can be worth the effort. Bad OCR severely impedes the workflow, especially if you can never trust whether the selected text is correct.

Be extremely careful with statistical or mathematical texts. If OCR misses a sign, e.g., a minus, the meaning changes completely (see image below, left the OCR text without minus signs, right the original).

# Digitizing Information Worksheet

To get information from books into a usable digital form, there are several options:

- **Using a Document Scanner:** The fastest option, but it destroys the book. See Box 4: Document Scanners and Box 5: Using a Document Scanner to Digitize Books.

## Box 3: Sufficient Quality Photos

Good photos of notes, or even book pages, mostly depend on good lighting. Current smartphone cameras are high-resolution enough for good results. Choose a place by a window with strong light, or use a lamp. Make sure the light comes from the side to avoid shadows from the smartphone or camera, and to reduce glare.

Document-scanner apps can improve the photo through post-processing. Alternatively, you can adjust exposure, contrast, white balance, and similar settings in photo-editing software, e.g., Photos on Mac, Affinity Photo, or Graphic-Converter. An «Unskew» function helps correct distortion, though the smartphone camera app may already show whether the phone is parallel to the surface via two crosses. If they are aligned, you are good. For most sketches, that is likely overkill.

If quality really matters, e.g., when digitizing an art book, consider a photo box and a prosumer-level camera with a tripod. A fixed camera allows you to use low ISO, e.g., 100 or 120, which means low camera noise, while avoiding camera shake when the shutter speed is too slow for handheld shooting.

- **Book Scanners:** These are too specialized for most private individuals, but may be available in public libraries (see Figure 2). You place the book on the scan unit and turn the pages while the scanner essen-



Figure 2: Book Scanner in a Library

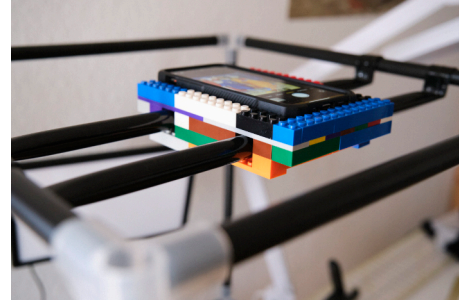
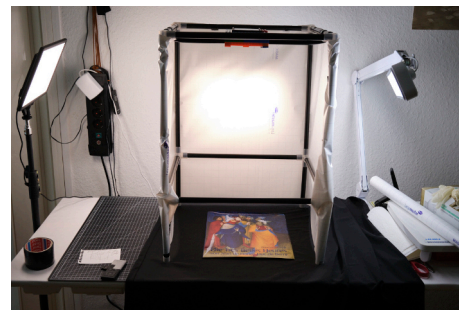


Figure 3: Photographing books with a DIY Photo box.

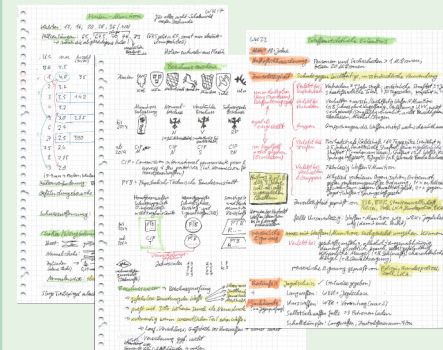
tially photographs the pages and processes them automatically, e.g., deskewing them. The result is usually a PDF.

- **Photographing Pages:** A smartphone with a photo or document-scanner app, or a more professional camera, can digitize the book while leaving it intact. A tripod or other stabilization, combined with a

## Box 4: Document Scanners

A document scanner has a page feed, meaning you can feed it about 50 pages and let it scan without further intervention. However, the pages must be separate. For books, that means cutting off the spine and essentially destroying the book (see Box 5: Using a Document Scanner to Digitize Books). If you scan sticky notes, putting them on a sheet of paper lets you scan six or more at once.

Good scanners have ultrasound overlap detection and stop when they try to scan two pages at once. This function will produce false alarms if you scan sticky notes on a sheet of paper, because they are, in fact, overlapping pages.



## DRAFT VERSION FOR FEEDBACK

remote, helps avoid camera shake and keeps the photos consistent. A photo box helps with lighting (see Figure 3).

- **Finding the Digital Version:** Check whether a legal digital version is available through a library, publisher, public-domain archive, or service you are allowed to use. You might find one online, e.g., at the

These scanners are not cheap, but they can quickly digitize documents such as contracts, certificates, insurance information, and similar material, allowing you to include them in your backups. They are also useful for digitizing handwritten study notes. Thus, they might be worth considering even if they would be too much just for paper idea notes.



# Digitizing Information Worksheet

DRAFT VERSION FOR FEEDBACK

## Box 5: Using a Document Scanner to Digitize Books

If you want to digitize paper books and are willing to destroy them, a document scanner can do it quickly. It also gives you clean scans without black borders around the page. You remove the spine so the book becomes a stack of separate pages the document scanner can handle.

A retractable carpet knife/cutter, a metal ruler, and a self-healing mat work best.

With hardcovers, remove the cover first. Open the book. If the first page is glued to the cover, remove that one first; otherwise cutting becomes difficult. Then cut through the paper connecting the cover to the pages until the cover is separated from the book. With the hardcover removed, you essentially have a paperback.



With paperbacks, place the metal ruler near the spine. Take care not to cut into the print. You need the right distance: far enough from the spine to separate the thread or glue binding, but not so far that you cut into the text. Then use the carpet knife to remove the

spine. With thicker books, first separating the book block into thinner blocks can help. Cutting is easiest with mass-market editions on cheap, «fluffy» paper, and hardest with thin paper.



What sounds gruesome for bibliophiles is not that difficult, as long as you restrict yourself to mass-market editions. Used books are particularly suited for this, since smells from previous owners, e.g., smokers, are not an issue in the digitized document. It can also be fairly quick: about 15 minutes per book. If you do a lot of them, you can digitize about 100 books on one weekend. At that volume, however, the bottleneck often becomes later use, not capture.

If you want to save the pages and the margins are wide enough, consider ring-binding them. There are tools to do this yourself. It can be useful for books you want available even during a long-term power outage, e.g., first-aid books. It is not pretty, but it works.

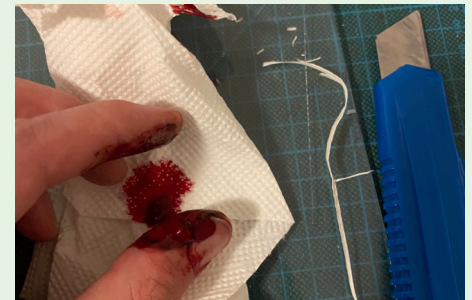
Destroy a book only if you accept the loss of the physical object, have a concrete reason to need the digital version, are legally allowed to

make the copy, and verify scan quality and backups before discarding the pages.

### Tips

- If you digitize many books, scan them by category, e.g., first all sewing books, then all photography books. Keep the most important category for last as motivation.
- Once you have tested the workflow on expendable books and know the scan quality is good, consider digitizing one emotionally difficult but replaceable book early. After that, less precious books become easier.
- Do not try to compete with the scanner while cutting books, e.g., cutting the next book while the previous one is scanned. The scanner will outrun you, and if you hurry, a book dies needlessly. First cut a stack, then scan, then OCR.
- Use scan profiles, e.g., Color-Duplex. For black-and-white books, such as most fiction books, scanning in black and white often produces better results. Test the options — color, greyscale, black-and-white, quality settings — and choose what suits you.
- For books that should be preserved in high quality, use at least 300 dpi and low compression.
- Keep a microfiber cloth ready to clean the scanner.
- Do not cut where you scan. It might make feeding the next stack easier, but cutting creates dust that gets into the scanner and produces scan artifacts.

- Use a metal ruler to guide the cutter. Plastic is a bad choice; the blade will cut into it. Do not force the blade or extend it too far. Otherwise the knife can jump or break. In the first case, you drive the blade into your thumb. In the second, a very sharp object flies uncontrollably in front of your face. Hopefully only in front of it.



- Elevate the blade more in the lower part of the book to cut deeper.
- Use a self-healing mat that does not blunt the knife once it is through the pages.
- After cutting, flip through the pages to ensure they are actually separated. Otherwise you will create a paper jam. Remove as much dust and glue residue as possible.
- Before throwing the paper pages away, make sure every page was scanned. Overlap detection is good, but you might still want to check manually or by calculating the page numbers.
- Take photos of hardcovers. Edit them — crop, contrast, white balance — and add them to the PDF.
- After scanning, use OCR (see Box 2: Optical Character Recognition (OCR)).

# Digitizing Information Worksheet

Internet Archive (<https://archive.org/details/books> or <https://archive.org/details/folkscanomy>), saving you the trouble of digitizing it yourself. See also Box 1: Rights Boundary.

- **Working with legally accessible digital books:** Sometimes you have legal access to a digital book and need excerpts, images, or pages available in your own working files. First use built-in export, citation, accessibility, or library functions. If screenshots are legally permitted, they can work well for short excerpts or selected pages. On a tablet, use the hardware buttons or define a shortcut. On a computer, use the screenshot shortcuts, ideally with the capture area limited to the book page. If the book can be shown as a single page and your display can be turned to portrait mode, you get a larger page view.

For repeated legally permitted captures, automation can reduce mechanical effort: page advance, screenshot, and filename by date/time can be handled with Automator on macOS; mass cropping with GraphicConverter; merging with Automator or a PDF app. Do not use this to bypass DRM, license terms, platform restrictions, or redistribution limits; see Box 1: Rights Boundary.

## Digitizing Whiteboards, Post-Its, etc.

Digital whiteboards exist, but analog whiteboards and sticky notes are often easier. No technical issues, except perhaps a dried-up marker. No distracting sound, no login problem, no digital options trying to seduce you

away from the work.

- **Photos:** Photos usually work best. If possible, turn the board to avoid glare. The first time, check the quality on a larger screen or zoom in to see whether even the smallest text is readable.
- **Post-It App:** If you use sticky notes and want to work with them digitally in a similar format, consider the Post-It App. It lets you photograph sticky notes, recognizes the individual notes (with manual assistance if needed), and can export them, e.g., into a PowerPoint presentation with individually movable notes (see Figure 4).

## Reduce File Size

Unless high quality is required, an A4-size image should be under 1 MB. If the file size is too large, you have a few options. Keep the originals on a backup drive. You never know whether you might later need the original quality.

- **Downscale Images:** You can reduce the dimensions, e.g., to 1200 pixels. If you do this often, consider automating it. Apps such as GraphicConverter allow you to set up AutoConvert, which monitors folders, automatically changes images placed inside them, and puts the changed image into your Collection Inbox (see Figure 5).
- **PDF Reduce Filesize:** Good PDF apps can reduce file size. Check the quality by zooming into the page and deciding what is still acceptable. Often it is easier to downscale the scanned images first and then turn them into a PDF.

## Trial Definition

If you want to improve how you digitize information, first assess what kind of material you have and what you actually need:

- What will this be used for?
- Does it need to be searchable?



Figure 4: Post-It App, detecting sticky notes.

## DRAFT VERSION FOR FEEDBACK

- Does the original need to be kept?
- Where will the digital version go?
- When will it be processed?
- What backlog limit stops the workflow?
- What are your hardware and software options, and what are your applications actually capable of?
- Where do you put what, and how do you name the files — destination folder/inbox, filename convention if needed, and when the item leaves the inbox?

Table 1 can guide the decision.

Explore different options and watch how much effort and time they take for the quality you get. Which options are sustainable? Once you find something that might work, do one trial. Do not redesign paper notes, book scanning, OCR, and audio transcription simultaneously.

For example, digitize this kind of information in this way for the next month. Afterwards, examine whether you actually did it and

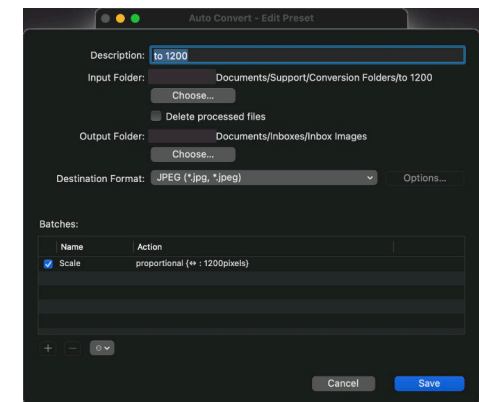


Figure 5: GraphicConverter Auto Convert

# Digitizing Information Worksheet

whether you actually used the information. A digitization workflow succeeds only if the material can be found and used later. Include at least one use-based success criterion, not only a count of created files.

Monitor possible failure modes:

- **Archive Spiral:** files increase, creative use does not.
- **Quality Trap:** rescanning/editing replaces using.
- **OCR False Confidence:** searchable text exists but contains errors.
- **Inbox Sedimentation:** captured scans never get processed.
- **Wrong Format:** beautiful image when selectable text was needed, or overbuilt PDF when a photo was enough.

## Examples

**Trial — handwritten idea notes:** For 3 weeks, photograph every paper idea note immediately after writing it and place it in the Collection Inbox folder. Each Friday, transcribe or discard all photographed notes.

- **Success:** At least 80% of photographed notes are processed within 7 days, and at least 5 are reused or linked to projects.
- **Abort:** Backlog exceeds 30 unprocessed photos or weekly processing takes more than 30 minutes twice.

**Trial — whiteboard capture:** For 4 weeks, photograph the whiteboard after every project session before erasing it.

- **Success:** All project sessions have a readable photo; at least 3 photos are later used to continue work.
- **Abort:** Photos are unreadable twice or not reviewed during the next session.

**Trial — book passages:** For 4 weeks, digitize only passages needed for active projects. Use phone scan + OCR or manual excerpt.

- **Success:** 10 passages added to project notes, and at least 5 are later used in drafting, outlining, or reference.
- **Abort:** More than 60 minutes/week spent cleaning scans, or more than 20 unused pages accumulate.

## Hand-Off

Having information available digitally — especially as text you can copy and paste — makes your Idea Collection much more flexible. Efficient search becomes possible, and changes become easier, especially when building on available information.

The options in this worksheet should help you find ways to digitize ideas, reference material, and documents, add them to your creative system, and create more efficiently.

Choose one material type and one workflow. Define a trial using the Integration Worksheet: trigger, start action, behavior, time source, success/abort criteria, trial duration, and review point.

Analog source	Minimum useful digital result	Usual best method	Trial metric
Handwritten idea	Searchable text or clear photo	Type, dictate, or photo + transcription	% captured notes usable after 7 days
Sketch	Legible image	Smartphone photo	Can reuse/reference it in project file
Whiteboard	Complete readable image	Photo	All important marks readable on laptop
Book passage	Selectable quote or page image	OCR / manual excerpt / scan	Can find and cite/use passage later
Whole book	Searchable PDF	Library scan/book or document scanner	Actually used in project within trial period
Audio note	Text transcript or indexed audio	Whisper/manual summary	Transcript good enough to reuse

Table 1: Selected Digitization Options