

scandata

JEFF THOMPSON + 2015

forward

The images in this book were scraped from the Internet Archive website. Used for calibrating large book scanners, they automate the process of color correction and the removal of hot spots created by the scanner's lights, and are meant to be read by the calibration software rather than humans.

When present in an Internet Archive entry, the images are stored in an odd file format (.ppm) inside a zip archive called "scandata.zip," from which the title for this book comes. Using a script written in Python, URLs for Internet Archive entries containing these files were pulled from the Bing search engine API, their contents downloaded, and the images converted to JPGs. This collection represents a small fraction of the thousands of such files on Internet Archive's site.

All images copyright their creators, everything else CC BY-NC-SA.

Details, 450 more images, and source code: jeffreythompson.org/scandata