

Thai OCR

OCR คืออะไร?



OCR นั้นย่อมาจาก **Optical Character Recognition** หรือ **Optical Character Reader (OCR)** ซึ่งเป็นกระบวนการของการแปลงสื่อสิ่งพิมพ์ เช่น กระดาษ นิตยสาร สัญญา หรือข้อมูลอะไรก็ตามที่อยู่ในรูปของเอกสารกระดาษหรือรูปภาพ ให้กลายเป็นข้อความโดยอัตโนมัติ ทำให้ไม่ต้องเสียเวลาพิมพ์งาน

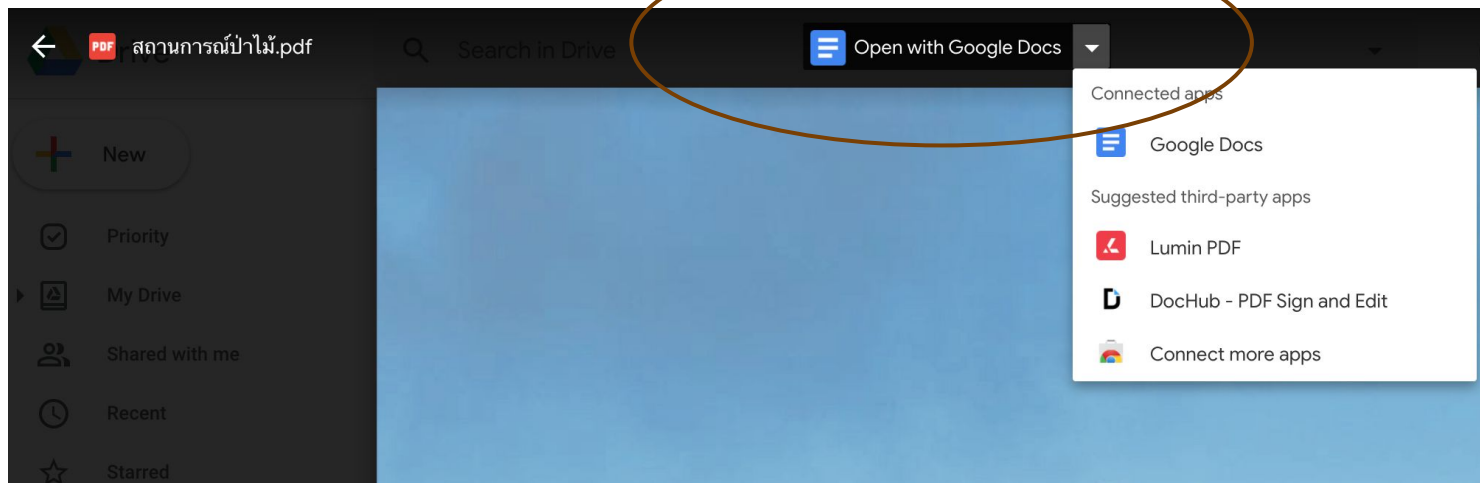
ปัจจุบันเทคโนโลยี **Optical Character Recognition** ได้ถูกทำให้ก้าวหน้ามาก แต่ก็ยังไม่ดีหากนำมาใช้งานกับข้อมูลประเภทลายมือหรือแบบอักษรที่มีลักษณะคล้ายกับลายมือ

ประโยชน์ที่ได้รับจากการใช้โปรแกรม OCR

- ประหยัดพื้นที่จัดเก็บข้อมูล เนื่องจากไฟล์ข้อความมีขนาดเล็กกว่าไฟล์ภาพมาก
- สะดวกในการปรับแต่งและแก้ไขเอกสาร เนื่องจากไฟล์ข้อความสามารถปรับแต่งและแก้ไขได้ง่ายกว่าไฟล์ภาพ

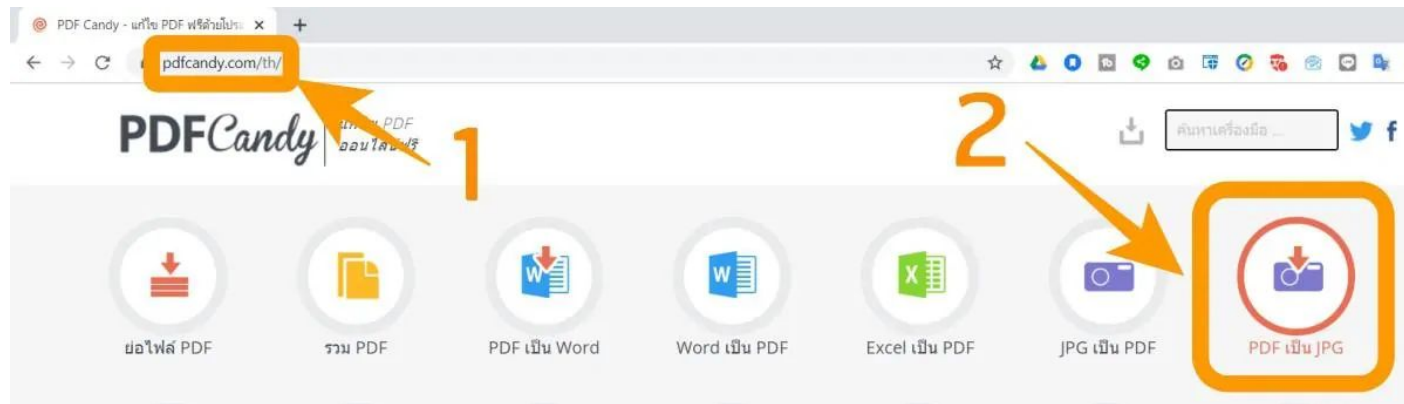
1. Google Doc

1. นำเข้า PDF file เข้า Google Drive เช่น [สถานการณ์ป่าไม้.pdf](#) หรือทดลองกับไฟล์ PDF ที่มี
2. เปิดไฟล์ และเลือก “Open with Google Doc”



2. PDFCandy >> <http://pdfcandy.com/th>

1. เริ่มจากเข้าเว็บไซต์ <http://pdfcandy.com/th>
2. แล้วคลิกที่ PDF เป็น JPG (ไม่ใช่ PDF เป็น WORD นะเดี่ยวแปลงไฟล์แล้วเพี้ยน ต้องเริ่มจากการแปลงไฟล์ PDF เป็น JPG ก่อน)



ที่มา: <https://www.it24hrs.com/2020/convert-pdf-to-word-for-thai-user/>

3. คลิกที่เพิ่มไฟล์ แล้วทำการอัปโหลดไฟล์ PDF ที่คุณเตรียมไว้แปลงเป็น WORD

PDFCandy | แก้ไข PDF | ออนไลน์ฟรี | แปลงจาก PDF | แปลงเป็น PDF | เครื่องมืออื่นๆ

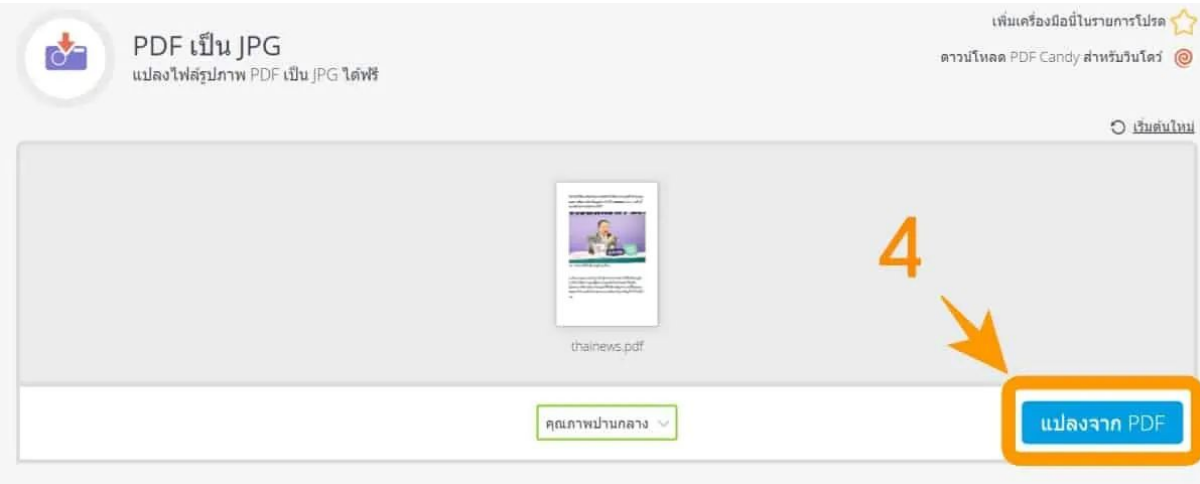
📄 🐦 f 🇹🇷 ภาษาไทย

PDF เป็น JPG
แปลงไฟล์รูปภาพ PDF เป็น JPG ได้ฟรี

เพิ่มเครื่องมือในรายการโปรด ☆
ดาวน์โหลด PDF Candy สำหรับวินโดวส์ 🎯

3 → + เพิ่มไฟล์ 
หรือลากและวาง

© ⓘ 🔄 🕒



4, หน้าจอแสดงตัวอย่างไฟล์ PDF ที่เราเตรียมไว้แปลงไฟล์เป็น WORD ซึ่งอัปโหลดลงบน PDFCANDY เรียบร้อยแล้ว ให้คลิกที่ แปลงจาก PDF ได้เลย

5. PDFCANDY จะแปลงไฟล์ PDF มาเป็นไฟล์ภาพ JPG ให้เรียบร้อย ให้คลิกที่ ดาวน์โหลดไฟล์





thainews.pdf

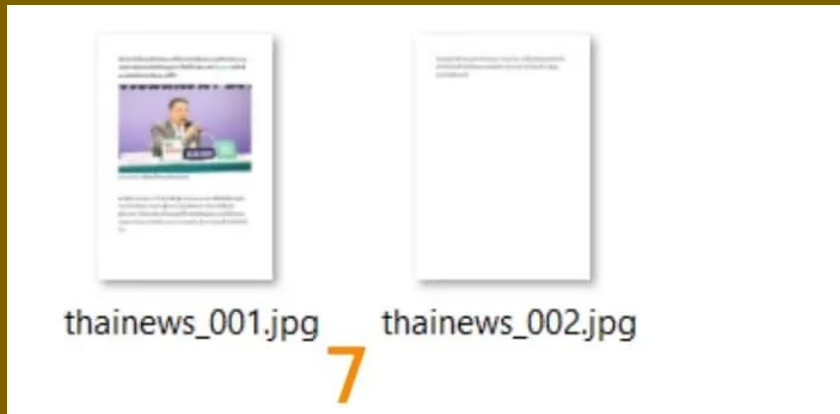


thainews.zip

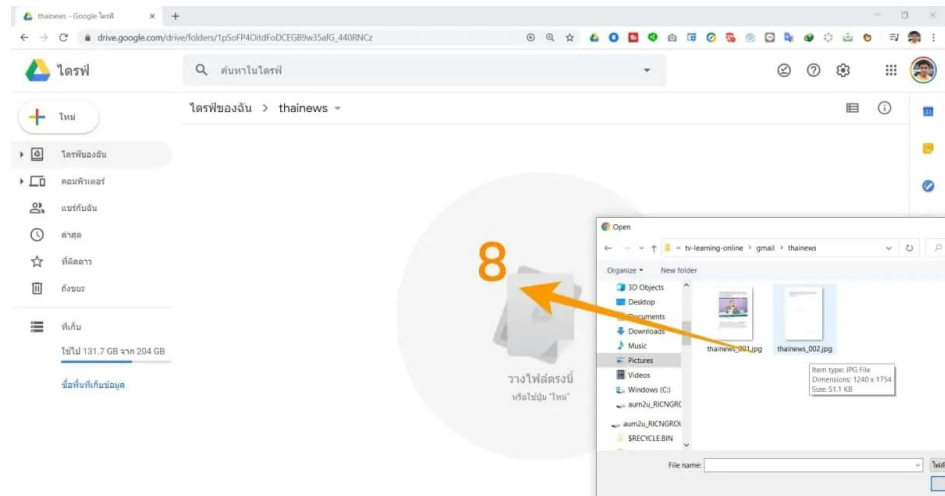


6. จะได้ไฟล์เป็น zipfile แบบนี้ ซึ่งในขั้นตอนต่อไปก็ให้คลิกขวาที่ตัวไฟล์แล้วเลือก **Extract all** เพื่อแตกไฟล์เป็นไฟล์ภาพ JPG เพื่อนำไปดำเนินการในขั้นตอนต่อไป

7. นี่คือตัวอย่างภายในไฟล์ PDF
แปลงไฟล์เป็นไฟล์รูป JPG
เรียบร้อยแล้ว

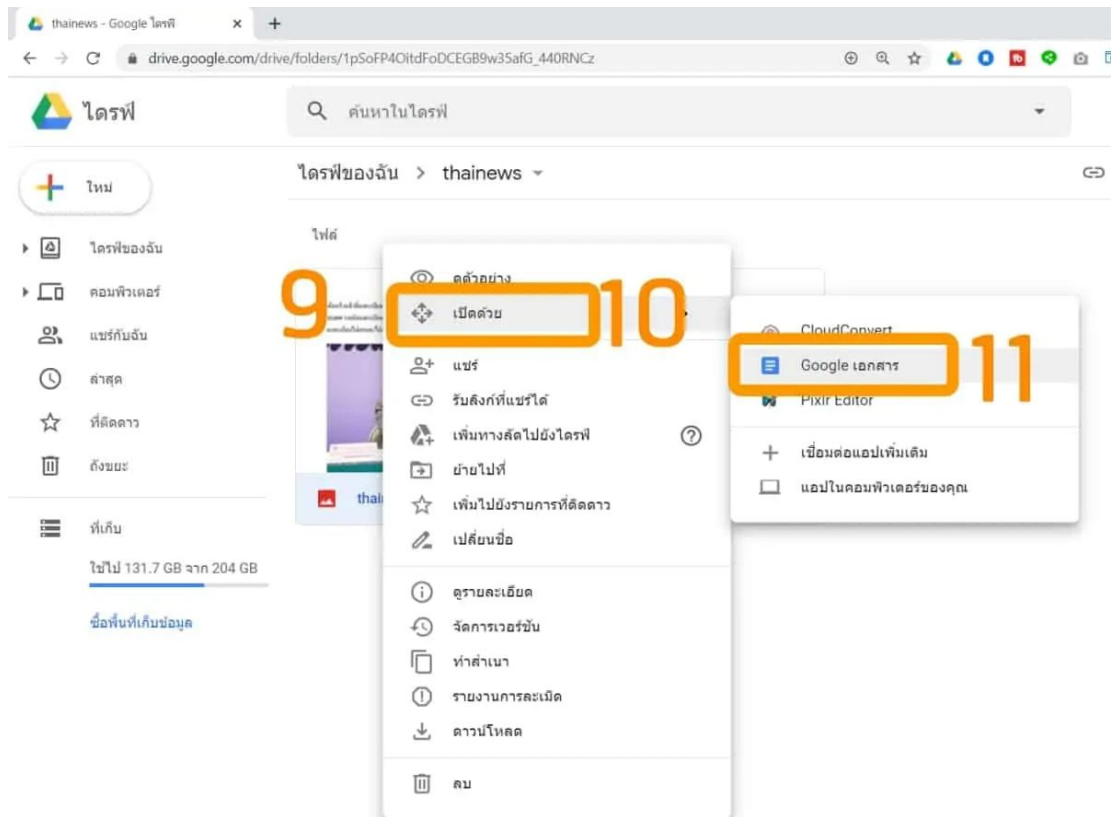


8. จากนั้นเข้า <http://drive.google.com> ซึ่งก็คือบริการ Google Drive นั่นเอง จากนั้นทำการ login ด้วยบัญชี Gmail ให้เรียบร้อยแล้ว นำไฟล์ภาพ JPG ที่แปลงจาก PDF เมื่อสักครู่นี้ อัปโหลดลงบน Google Drive เพื่อดำเนินการในขั้นต่อไป



9-11. เมื่ออัปโหลดไฟล์ภาพ JPG เรียบร้อยแล้ว ให้คลิกขวาที่ไฟล์ภาพ >> เลือก เปิดด้วย >> เลือก Google เอกสาร

ระบบทำการเปิดไฟล์ใน Google Docs โดยจะปรากฏรูปภาพในหน้าแรก และข้อความในหน้าถัดไป



3. PDF2JPG >> <https://pdf2jpg.net/>

ขั้นตอนจะเหมือนกับโปรแกรมอื่น คือ แปลง PDF ให้เป็นรูปภาพก่อน และนำเข้าสู่ Google Drive เพื่อเลือกเปิดด้วย Google Doc

Pick the PDF document to convert to JPG

The PDF size is limited to 25MB.

Choose a PDF file

No PDF selected

JPG quality

"Good" is the right choice for most PDF

Good - 150dpi - For PDF with illustrations



It's almost done! You're a click away from the PDF to JPG conversion

Convert PDF to JPG

4. I2OCR >> <http://www.i2ocr.com/free-online-thai-ocr>

1. เลือก Image OCR หรือ PDF OCR และกดเลือกไฟล์ที่ต้องการ ตามขั้นตอน และกด Extract Text เพื่อประมวลผล Image >> Text หรือ PDF >> Image

The screenshot displays the i2ocr.com website interface. At the top, there are two tabs: "Image OCR" (selected) and "PDF OCR" (marked as "NEW"). Below the tabs, the interface is divided into three steps:

- Step 1 Select Language:** Shows four service icons: "FREE EDIT PDF", "FREE CONVERT PDF TO WORD", "FREE PSD FILLER", and "FREE CONVERT PDF TO JPG". Below these is the "PDFSimpli" logo and a language dropdown menu set to "Thai".
- Step 2 Select Image:** Features radio buttons for "File" (selected) and "URL", and a "Select Image" button.
- Step 3 Extract Text:** Includes a "Select Checkbox" section with an unchecked box for "I'm not a robot" and a reCAPTCHA logo. A large red "Extract Text" button is positioned at the bottom right.

A large brown arrow points from the "Extract Text" button towards the right side of the image.

Thank You



This lesson was adapted from the World Bank's Introduction to Data Literacy training manual by Eva Constantaras, and adapted by Yan Naung Oak, Open Development Cambodia and Open Development Initiative, and is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). For full terms of use, see [here](#).