UNIDIR Cartridge Identification Planning

Below are some of the main AI/vision problems we need to solve.

In addition, we need to gather requirements for app capabilities besides the vision problems.

Finally, to reach the point where Microsoft can make a more significant investment, we need to be “data ready”, which entails building the capacity to collect and label a significant amount of data- ie thousands of cartridge images.

Mini deck with some sample images: <https://microsoft.sharepoint.com/:p:/t/CartridgeOCRforUNIDIR/ETcurCw1QDNBgbE_3WojtYIBawc3fNKpLAqinK_aTMRsyA?e=oIdrPy>

Core AI problems:

* Locate cartridge in image
  + Assuming well-focused good resolution still frame
  + [stretch] from video frames with motion, variable focus.
* Headstamp OCR
  + By “unrolling” and automatically locating the text
  + Using raw image -> text, eg with a neural net.
  + [stretch] Non-Latin character sets.
* Phone app UX features
  + Snap image and auto-extract
  + Geolocation or manual entry
  + Other site metadata [ get specs from David ]

Locating cartridge:

Input: still frame

Output: center and radius of cartridge

Options:

* Take a purely geometric approach, eg using Hough transform.
* Use a convolutional network.
  + Requires training data. Do we need a tool for manually labeling images? This might be useful for other purposes later.
  + We might be able to do a good job with a relatively small number of images- a few hundred?

OCR:

Ideally we can just unroll or warp the image to produce straight-line text and apply off-the-shelf OCR. I suspect this will yield poor quality for these images. Maybe there are some enhancements we can do to the image, or we can take advantage of multiple video frames, to produce a higher-quality OCR input.