MEETING NOTES
Signs of Literacy Kaggle Steering Committee
Wednesday, April 3rd 2019

ATTENDEES:

Mark Hailwood (Bristol), Judy Stephenson (UCL), David Underdown (TNA), Josh Ivinson (Cambridge), Colin Greenstreet (MarineLives)

APOLOGIES: Sara Caputo (Cambridge); Michael Bennett (Sheffield); Richard Blakemore (Reading)

DOCUMENTS:

The link to the slide pack we reviewed is here:

https://github.com/Signsofliteracy/ArchiveBots/blob/master/SOLM_SC_Slides_FINAL_03042019.pdf

ACTIONS:

1. Revisions to Signs of Literacy initiatives objectives requested from all to incorporate broader vision of what we aim to do as a group (see current version below). RESPONSIBLE Mark, Judy, David, Josh, Michael, Sara [By end April 2019]

2. Finalise objectives, desired outputs and rules for first Kaggle competition and agree these with Kaggle/Google. Colin to draft short note suitable for use with Kagglers as well as Kaggle/Google. The note will be along lines of today's discussion: (a) Kagglers to identify which images contain (i) signoffs (mark(s), initial(s), signature(s) and to place a boundary box round one or more signoffs on a page (b) Kagglers to identify which images contain "metadata" (defined as Deponent name, deponent occupation, deponent age, deponent place of residence, date of deposition) and to place a boundary box around one or more metadata areas on a page (c) Kagglers to identify the correct match between identified metadata areas and identified signoff areas. RESPONSIBLE: Colin [By end April 2019]

3. Colin to check with John Sheridan, TNA digital director, whether any further signoff is required on image use prior to Colin signing contact with Google on behalf of Chronoscopic Education (which will be the contracting party with Google). RESPONSIBLE: Colin [By end April]

4. Agreed that the next SOLM data set [SOLM-2019] will consist solely of English High Court of Admiralty deposition images. It is estimated it will comprise ca. 100,000 images and will be fully imaged (but not fully annotated or otherwise indexed) by end calendar 2019. SOLM-2018 contains ca. 40,000 English High Court of Admiralty images, so we will be scaling up 2.5x. RESPONSIBLE: Colin [By end December 2019]

5. Agreed for Colin to sit down with Josh and Michael in ca. 3 months to discuss and structure a potential TNA + Josh + Michael small scale project with TNA as the PI and funding bidder to run from (say) 3Q 2020 to test machine learning tools we will by then have developed for the ability to enhance Josh and Michael's research productivity. Colin to bring Michael up to speed on this idea and to check his interest. Josh has already signalled interest. RESPONSIBLE: Colin, Josh, Michael [By end August]

6. Agreed to set up meeting with Pip Willcox, Head of Research, UK National Archives, (former Curator of Digital Special Collections, Bodleian Library) [https://twitter.com/pipwillcox?lang=en] to (a) Understand Pip's objectives for the TNA (b) Propose a fundedresearch collaboration between TNA + Signs of Literacy in 2H 2020 and beyond, to test the impact of machine learning tools on research productivity, discoverability and accessibility using the SOLM-2019 data set, comprising a full set of images of all HCA depositions between c. 1570 and ca. 1680 [HCA 13/20 - HCA 13/76]. RESPONSIBLE: Colin, David; with potential participation at a follow up meeting by Josh and Michael [Target meeting September 2019]

7. Explore potential image sourcing and volunteer participation partnerships with (a) UK county and municipal archives (b) Spanish and French notarial archives. RESPONSIBLE: Colin + potentially Zoe Screti (PhD candidate, University of Birmingham). TIMING: 2H 2019; 1H 2020. Low key.

8. Identify a network coordinator and initiate a network of historians (and non-historians) interested in historical literacy and in the Signs of literacy initiative. Ask Josh, Michael and Sarah whether this might interest one of them. Mark Hailwood has offered to discuss the functioning of such a network with any one who wishes to explore this role, drawing on his experience of several networks he has coordinated. RESPONSIBLE: Mark, Josh, Michael, Sarah [By end May for decision on coordinator; timing of development of network for discussion once a network coordinator has been identified and agreed]

I hope this captures the actions resulting from today's meeting.

Do please make corrections and additions. I am heading to Majorca for a few days, back on Sunday, but will be reading my emails while enjoying (I hope) the Majorcan sun and catching up with a friend who has recently moved there.