

# Revisiting historical literacy

DHC 2018

Thursday, September 6<sup>th</sup>, 2018  
University of Sheffield

Dr Mark Hailwood  
Colin Greenstreet

# Digital Humanities and Early Modern History

The screenshot shows the homepage of the English Broadside Ballad Archive. At the top, there's a banner with the text "UNIVERSITY OF CALIFORNIA SANTA BARBARA English Broadside Ballad ARCHIVE". Below the banner, a red bar contains a search field labeled "BALLAD SEARCH" with "Advanced Search" and "Search" buttons. To the right of the search bar is a large image showing various historical broadside ballads and musical notation. The main content area has a white background with a heading "Making broadside ballads of the seventeenth century fully accessible as texts, art, music, and cultural records." Below this, a sub-headline says "Click below to explore sample topics in early modern popular culture:". There are four thumbnail images: "House of Cards" (Early Modern Political Intrigue), "The Bachelor" (Early Modern Love), "KENTISH DICK" (Early Modern Crime), and another broadside. On the left, a sidebar titled "ARCHIVE AREAS" lists links to Home, About Us, Features, Holdings, Resources, Use Policy, Visualizations, Contact Us, and Early Modern Center.

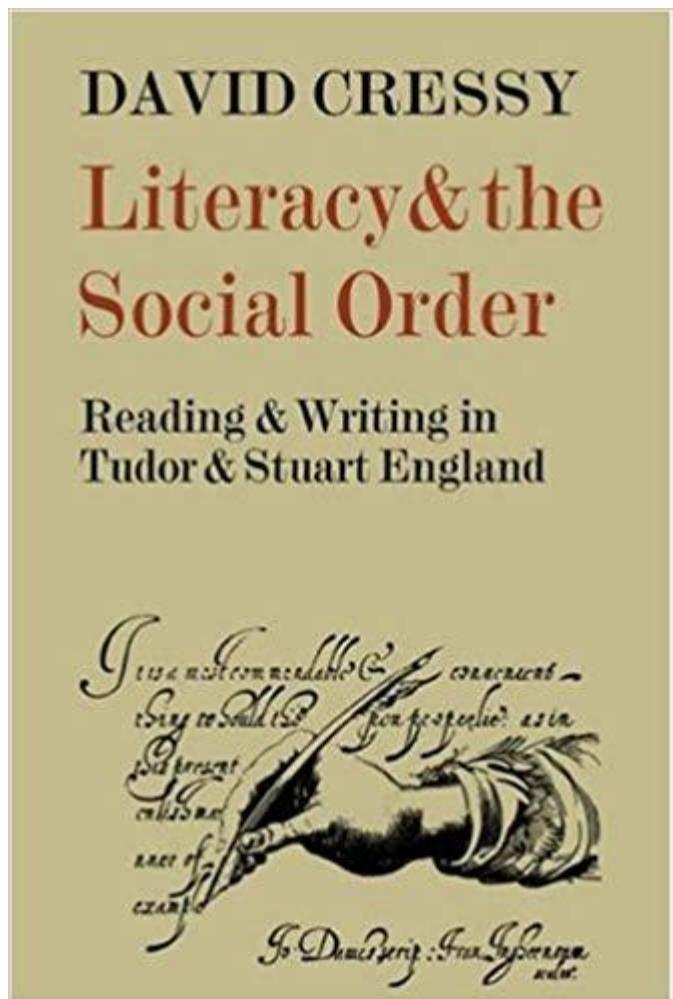


- [ABOUT EEBO](#)
- [EEBO DEMO](#)
- [EEBO FULL TEXT DEMO](#)
- [REQUEST A TRIAL](#)
- [CONTACT](#)



From the first book printed in English by William Caxton, through the age of Spenser and Shakespeare and the tumult of the English Civil War, Early English Books Online (EEBO) will contain over 125,000 titles listed in Pollard and Redgrave's Short-Title Catalogue (1475-1640), Wing's Short-Title Catalogue (1641-1700), the Thomason Tracts (1640-1661), and the Early English Tract Supplement - all in full digital facsimile from the Early English Books microfilm collection.

Who could read?



1980

Table 1: Levels of illiteracy in early modern England  
(Cressy, 1980)

Date	Men	Women
1485–1507	90%	99%
1558	80%	95%
1642–1645	70%	90%
1714	55%	75%

**the many-headed monster**

the history of 'the unruly sort of clownish' and other early modern peculiarities

Search



About Us   Browse By Theme   Friends   Monster Mini-Series   History from Below   Voices of the People

Posted on October 13, 2014 by manyheadedmonster

— Previous   Next —

 **'The Rabble that Cannot Read'?  
Ordinary People's Literacy in  
Seventeenth-Century England**

Mark Hallwood

Those of us historians intent on exploring the world of ordinary women and men in the sixteenth and seventeenth centuries conduct a lot of our research by looking at surviving examples of what such people read—for instance, cheap printed broadside ballads—or of what they wrote—take, say, Joseph Button's notebooks. These materials are fascinating and undoubtedly useful, but regular readers of this blog might understandably find themselves wondering about the validity of this approach, and asking themselves a simple but important question: to what extent could the lower classes of England actually read and write in the seventeenth century?



— David Teniers the Younger 'Peasant Reading a Letter...' But could they?

Source:

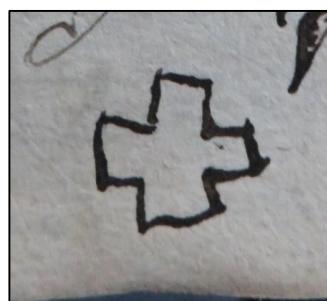
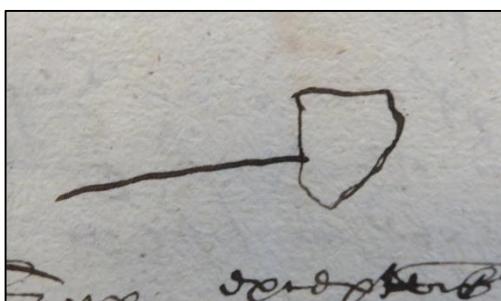
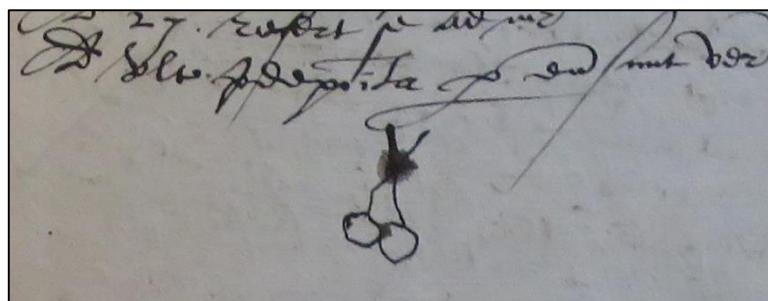
<https://manyheadedmonster.wordpress.com/2014/10/13/the-rabble-that-cannot-read-ordinary-peoples-literacy-in-seventeenth-century-england/>

Enoch Bufford  
John Hedges

Widow of spider  
her mark

William H Hickman  
his mark

Richard El Guy  
his mark



John Proporto  
G. J. G. Libbott  
Aug 16th 1888

Carroll, Ohio  
Canton Evans  
is

Proctor  
Fogel 2nd October 1

## Reading, Writing, and Initialing: Female Literacy in Early Modern London

Eleanor Hubbard

**Abstract** This article reopens the vexed question of how many women in early modern England could read by calling attention to the precise ways in which women marked, initialed, and signed legal depositions in late sixteenth- and early seventeenth-century London. It shows that initialing and signing were closely correlated skills, and it argues that women who wrote their initials had begun to learn how to read. Using initials as a proxy for elementary reading literacy, it goes on to map female literacy in early modern London, showing that urban upbringings fostered female literacy and that reading literacy was far more broadly socially diffused than the ability to write. Changes in initialing patterns as women aged suggest that women found reading to be useful and relevant to their lives, and that literacy carried social prestige.

In 1611, a cooper's wife named Mary Swainie told the London Consistory Court that she got her living "by her husbands trade & by teaching of children."<sup>1</sup> She was, it seems, one of the innumerable informal schoolmistresses who taught children to read in early modern England. Tacitly tolerated by the ecclesiastical visitors who were supposed to license all teachers, they usually entered the historical record only when they caused offense, or when they were criticized by professional schoolmasters for their ignorance.<sup>2</sup> Indeed, Mary Swainie, whose profession was recorded only because she happened to depose in a testamentary case, was not a well-educated woman. Instead of signing her deposition, she marked it with a clear initial *M*—the first letter of her first name (see Figure 1). A historian of literacy counting signatures would typically mark her as illiterate and perhaps



**Signs of Literacy ||**

A community project for collaborative research into historical literacy

# Some perspective

**Labeled Faces in the Wild**

UNIVERSITY OF MASSACHUSETTS AMHERST, MASS.

Labeled Faces in the Wild Home

**Menu**

- LFW Home
  - Mailing
  - Explore
  - Download
  - Train/Test
  - Results
  - Information
  - Errata
  - Reference
  - Resources
  - Contact
  - Support
  - Changes
- Part Labels
- UMass Vision

**NEW SURVEY PAPER:**

Erik Learned-Miller, Gary B. Huang, Aruni RoyChowdhury, Haoxiang Li, and Gang Hua.  
**Labeled Faces in the Wild: A Survey.**  
In *Advances in Face Detection and Facial Image Analysis*, edited by Michal Kawulok, M. Emre Celebi, and Bogdan Smolka, Springer, pages 189-248, 2016.  
[Springer Page] [Draft pdf]

**NEW RESULTS PAGE:**

WE HAVE RECENTLY UPDATED AND CHANGED THE FORMAT AND CONTENT OF OUR [RESULTS PAGE](#). PLEASE REFER TO THE [NEW TECHNICAL REPORT](#) FOR DETAILS OF THE CHANGES.

Welcome to Labeled Faces in the Wild, a database of face photographs designed for studying the problem of unconstrained face recognition. The data set contains more than 13,000 images of faces collected from the web. Each face has been labeled with the name of the person pictured. 1680 of the people pictured have two or more distinct photos in the data set. The only constraint on these faces is that they were detected by the Viola-Jones face detector. More details can be found in the technical report below.

There are now four different sets of LFW images including the original and three different types of "aligned" images. The aligned images include "funneled images" (ICCV 2007), LFW-a, which uses an unpublished method of alignment, and "deep funneled" images (NIPS 2012). Among these, LFW-a and the deep funneled images produce superior results for most face verification algorithms over the original images and over the funneled images (ICCV 2007).

**Related:**

[new] Collected resources related to LFW - updated 2017/05/09.  
LFW Deep Funneled Images.  
LFW attributes file (see Attribute and Simile Classifiers for Face Verification, Kumar et al.).  
Face Detection Data set and Benchmark (FDDB), our new database for face detection research.  
Faces in Real-Life Images workshop at the European Conference on Computer Vision 2008, run by Erik Learned-Miller, Andras Ferencz, and Frederic Jurie.

**Abstract** In 2007, Labeled Faces in the Wild was released in an effort to spur research in face recognition, specifically for the problem of face verification with unconstrained images. Since that time, more than 50 papers have been published that improve upon this benchmark in some respect. A remarkably wide variety of innovative methods have been developed to overcome the challenges presented in this database. As performance on some aspects of the benchmark approaches 100% accuracy, it seems appropriate to review this progress, derive what general principles we can from these works, and identify key future challenges in face recognition. In this survey, we review the contributions to LFW for which the authors have provided results to the curators (results found on the LFW results web page). We also review the cross cutting topic of alignment and how it is used in various methods. We end with a brief discussion of recent databases designed to challenge the next generation of face recognition algorithms.

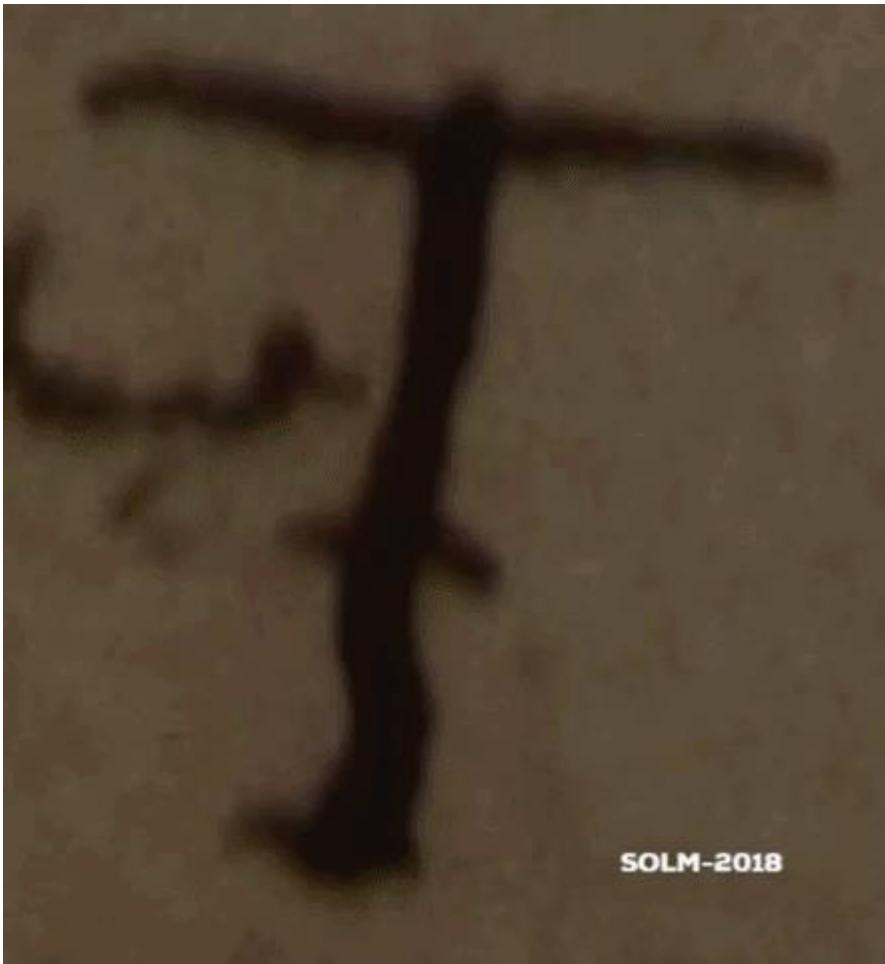
## Labeled Faces in the Wild: A Survey

**Erik Learned-Miller, Gary Huang, Aruni RoyChowdhury, Haoxiang Li, Gang Hua**

In *Advances in Face Detection and Facial Image Analysis*, edited by Michal Kawulok, M. Emre Celebi, and Bogdan Smolka, Springer, pages 189-248, 2016.

# Pattern seeking

How many different letters can you recognise?



31 images of initials in the SOLM-2018 database  
Animated GIF viewable only in Powerpoint

# Initials – Ts and Js, and a few others thrown in



# Our vision is a SOLM-2023 database with 1 million marks, initials & signatures from across Europe & North America from the C16th to C18th

## The maths

- 3 person/months to create 5,000 signoff SOLM-2018 database consisting of image snippets; boundary boxed snippets on full page images; 5,000 lines x 25 rows of metadata
- 6 person/months to create our targeted 10,000 SOLM-2018 training database
- 20,000 signoff processing per person year
- Target of 1 million signoffs in our database
- 100,000 signoffs per year with 5 people working full time

That's TEN YEARS to achieve our vision  
with 50 person years to do it!!!!

The **SOLM-2018 database** is a tool for historians and computer scientists to work with marks, initials and signatures. It has been designed to support the exploration of historical literacy and the development of tools for automatic metadata creation.

We will be previewing the database at the TNA Archives & AI symposium on Tuesday, September 4<sup>th</sup> and at the Sheffield Digital Humanities Congress on Thursday, September 6<sup>th</sup>, 2018.

We are looking for UK and international archival partners interested in contributing content to the SOLM-2018 tool and in learning about AI based pattern recognition.

We are especially interested in manuscripts containing marks, initials and signatures by individuals engaged in marine and shore trade occupations from the following English towns and areas for the C16th, C17th and C18th:

Aldeburgh [Suffolk]	Dover [Kent]	Ipswich [Suffolk]	Scarborough [North Yorks]
Barnstaple [Devon]	Falmouth [Devon]	Newcastle	Weymouth [Dorset]
Bermonsey	Faversham [Kent]	Plymouth [Devon]	Woodbridge [Suffolk]
Bristol	Foy [Cornwall]	Rochester [Kent]	Yarmouth [North Yorks]
Colchester [Essex]	Greenwich	Rotherhithe	
Dartmouth [Devon]	Harwich [Essex]	Southampton	
Deptford	Hull	Southwark	

For further information contact Colin Greenstreet, community organiser, Signs of Literacy initiative, or Dr Mark Hailwood (Bristol)  
GitHub: <https://github.com/Signsofliteracy>

5000 signoffs  
and growing

**Our challenge to archivists, computer scientists and historians:** Help us develop the tools to create a SOLM-2023 database of 1 mill signoffs with a productivity rate of ten times today's best, at a resource cost of 5 person/years, not 50 person/years, and in half the time

More generally, we need to work together, if we are going to make sense of our digitised manuscript archives – **developing AI tools to process archival images and to identify, extract, read and record metadata**

For more information contact Colin Greenstreet, community organiser of the Signs of Literacy initiative, and Dr Mark Hailwood (Bristol)  
<https://github.com/Signsofliteracy>

# Digitisation of Early Modern manuscripts, machine learning & collaborative working will enable great advances in understanding of the granularity of historical literacy

Cressy (1980) sample sizes in brown; SOLM-2018 (n=5,000) sample sizes in green

SOLM-2023 goal (n = 1 mill)

## 134 *Literacy and the social order*

Table 6.7 *Ranking of London and Middlesex trades by illiteracy, 1580–1700*

Trade	No. sampled	No. mark	% mark
Scrivener	46	0	0
<b>Merchant</b>	<b>27</b>	<b>0</b>	<b>0</b>
Vintner	23	0	0
Grocer	21	0	0
Saddler	10	0	0
Apothecary	9	0	0
Goldsmith	29	1	3
Stationer	18	1	6
Chandler	28	2	7
Barber	13	1	8
Ironmonger	13	1	8
Draper	34	4	12
Haberdasher	49	7	14
Dyer	13	2	15
Glazier	13	2	15
Leatherseller	12	2	17
Skinner	12	2	17
Cutler	16	3	19
<b>Cooper</b>	<b>13</b>	<b>3</b>	<b>23</b>
<b>Mariner</b>	<b>13</b>	<b>3</b>	<b>23</b>

27 | 875

13 | 72

13 | 2348



Brown = Cressy (1980)

Green = SOLM-2018 (2018, half complete)

## The structure of illiteracy

## 135

Table 6.7 cont.

Trade	No. sampled	No. mark	% mark
Baker	19	5	26
<b>Cook</b>	<b>11</b>	<b>3</b>	<b>27</b>
Victualler	40	11	28
<b>Sailor</b>	<b>17</b>	<b>5</b>	<b>29</b>
Clothworker	50	9	30
Merchant-Taylor	62	19	31
Cordwainer	42	13	31
Weaver	29	10	34
Butcher	26	9	35
Blacksmith	37	14	38
Joiner	32	12	38
Bricklayer	21	8	38
<b>Carpenter</b>	<b>40</b>	<b>16</b>	<b>40</b>
Feltmaker	10	4	40
Innkeeper	10	4	40
Tailor	97	43	44
Brewer	13	6	46
Shoemaker	19	10	53
Gardener	17	10	59
Waterman	15	10	67

11 | 32

17 | 74

40 | 128

15 | 71

# Signs of Literacy Kaggle Research Competition

Signs of Literacy Kaggle Research Competition, 2018  
Colin Greenstreet on LinkedIn  
April 30, 2018

Google owned Kaggle has selected us as one of a small number of pro bono competitions they support each year on the merits of our proposal, and the potential impact on the research field and community of the competition.

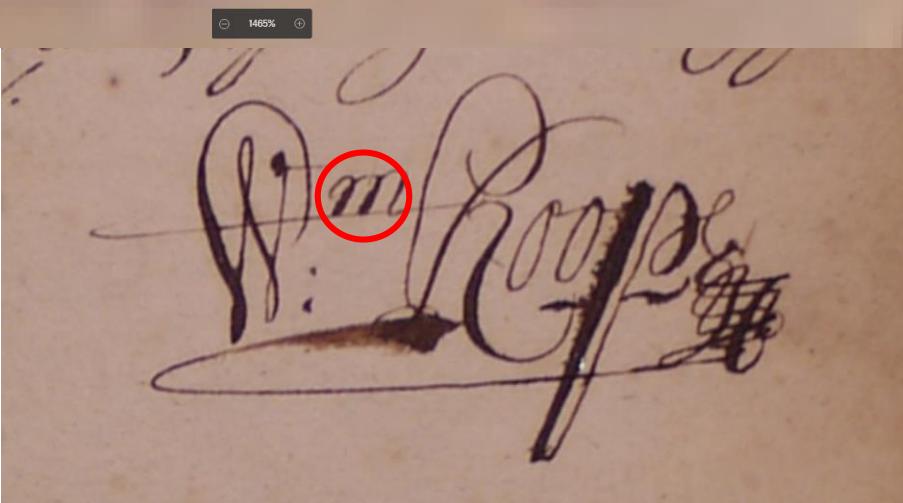
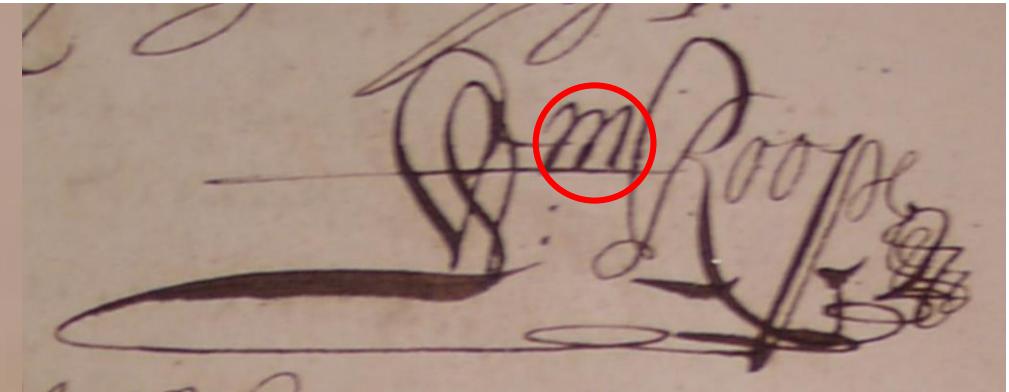
Kaggle will cover the running costs of the competition. We will provide the prize pool, and are now seeking to raise US \$30,000 from potential sponsors and partners.

The Proof of Concept will contain two parts:

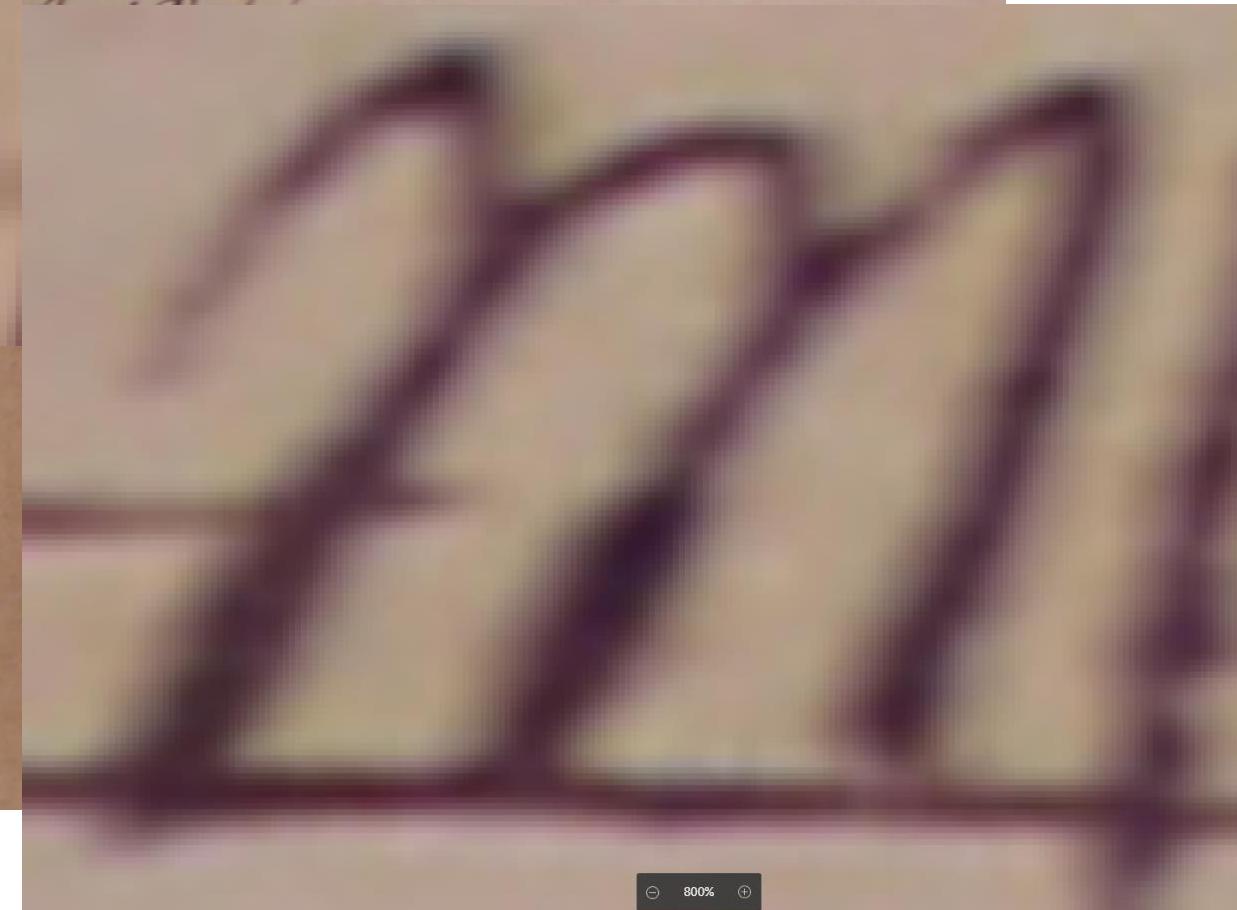
- (1) Algorithmic identification of marks, initials and signatures.
- (2) Algorithmic discrimination between degrees of "sophistication" within the three categories of "mark"; "initial(s)", and "signature".

Having proven the concept, we will seek out an image or vision oriented computational laboratory with which to develop a grant funded collaboration to take the work further in 2019 and beyond.

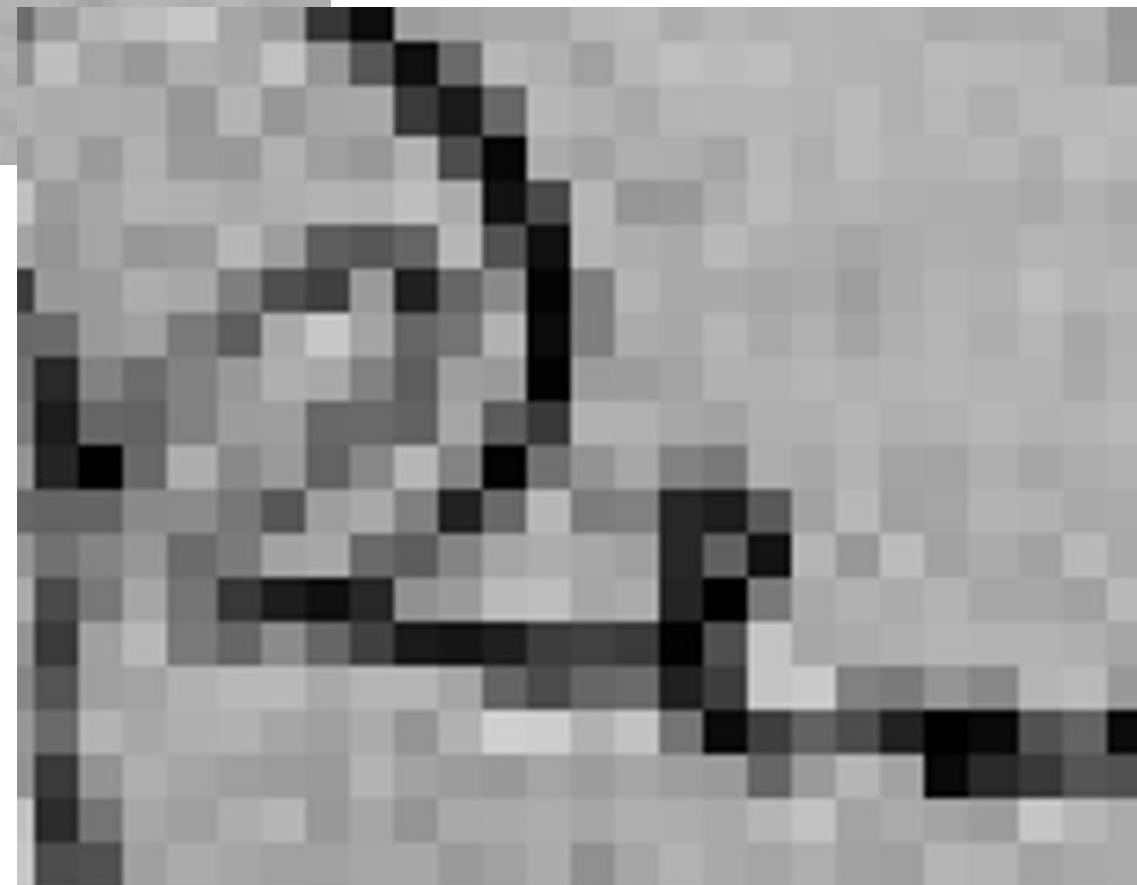
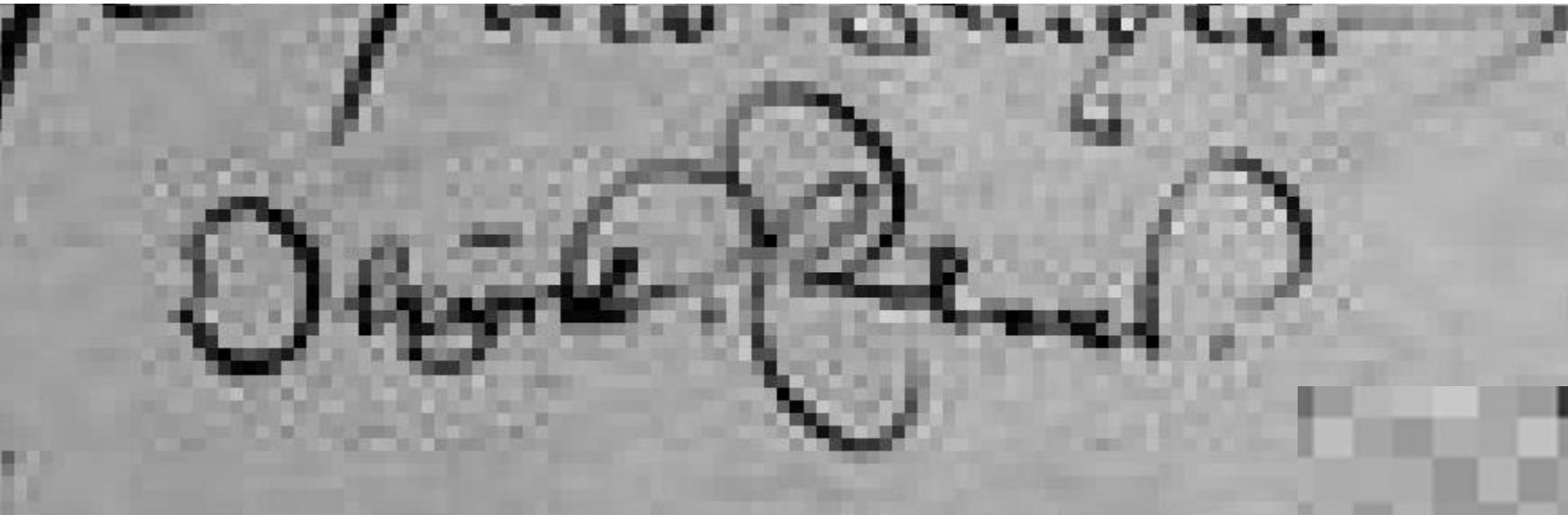
# High pixel definition



Source: KaggleTestSnippet\_HCA\_1373\_f.199r.PNG,  
KaggleTestSnippet\_HCA\_1373\_f.199v\_One.PNG



## Low pixel definition



# Colour analysis – image colour extract PHP, hexadecimal colours

The figure consists of four panels, each showing a snippet of handwritten text and its corresponding color analysis results.

- Panel 1:** Shows the text "Nicholas Harrison". The color analysis table is as follows:

Color	Color Code	Percentage
#e0a080	0.855975	
#c08060	0.084403	
#806040	0.039371	
#604020	0.013208	
#402020	0.007044	

- Panel 2:** Shows the text "Bo:nglis". The color analysis table is as follows:

Color	Color Code	Percentage
#c08060	0.883721	
#806040	0.063798	
#604020	0.048605	
#402000	0.003876	

- Panel 3:** Shows the text "Jacob pintorB". The color analysis table is as follows:

Color	Color Code	Percentage
#808080	0.969271	
#604040	0.030729	

- Panel 4:** Shows the text "H. Langius". The color analysis table is as follows:

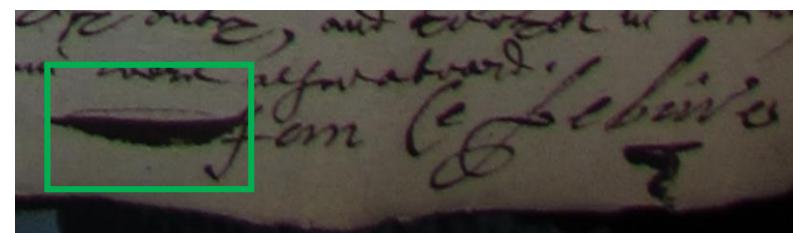
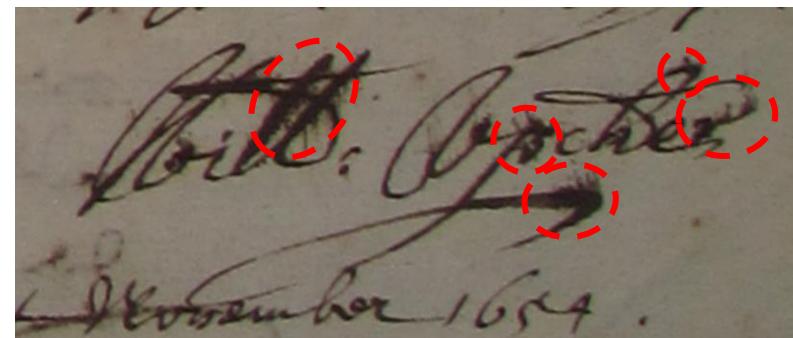
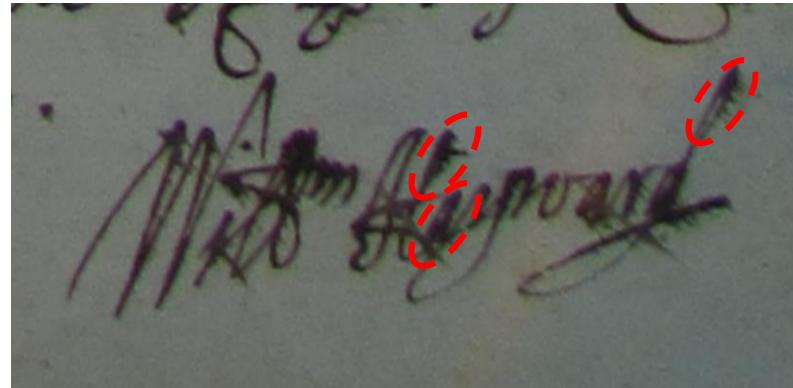
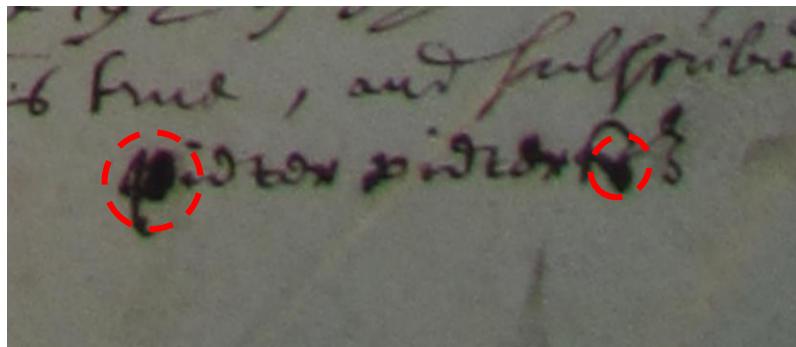
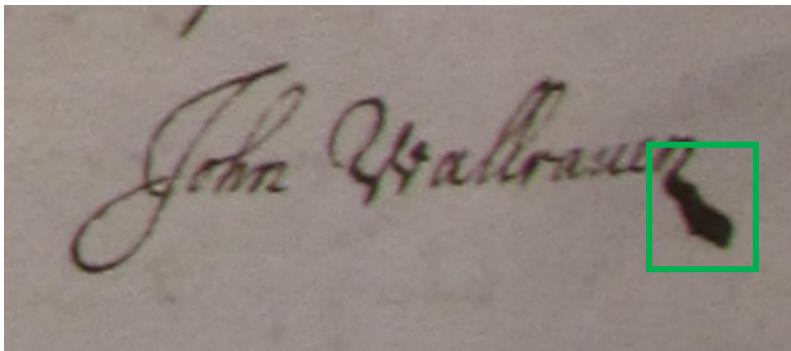
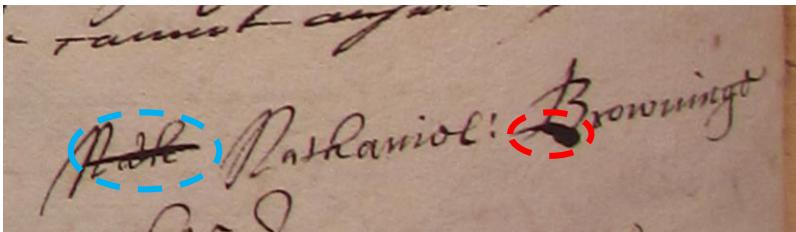
Color	Color Code	Percentage
#806040	0.806762	
#604020	0.125143	
#402020	0.062000	
#202000	0.006095	

Below each panel is a grayscale version of the same image, indicating the processing steps taken.

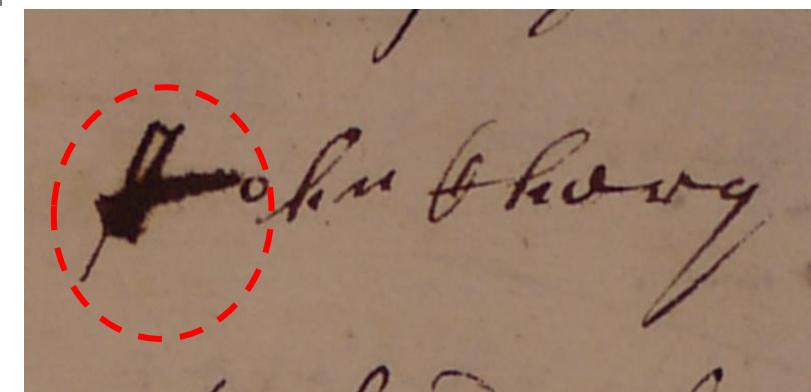
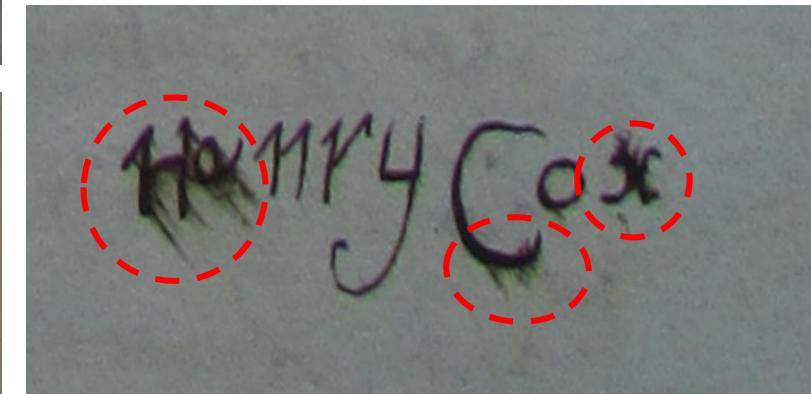
Source: Sample images from SOLM-2018 (KaggleTestSnippet\_HCA\_1353\_f.275v.PNG, KaggleTestSnippet\_HCA\_1353\_f.270v\_Two.PNG,

KaggleTestSnippet\_HCA\_1370\_f.463r\_One.PNG, KaggleTestSnippet\_HCA\_1368\_f.497v.PNG) processed in [http://www.coolphptools.com/color\\_extract#demo](http://www.coolphptools.com/color_extract#demo); same images reprocessed in Photos SW package, with adjustments set to 0% light, 0% colour, 100% clarity

# Detection and analysis of blots, smudges, stylistic features, & deletions



Source: Clockwise from top LH side:  
KaggleTestSnippet\_HCA\_1370\_f.387v.PNG,  
KaggleTestSnippet\_HCA\_1370\_f.13r.PNG,  
KaggleTestSnippet\_HCA\_1370\_f.167r.PNG,  
KaggleTestSnippet\_HCA\_1371\_f.456r.PNG,  
KaggleTestSnippet\_HCA\_1370\_f.15r.PNG,  
KaggleTestSnippet\_HCA\_1370\_f.19r.PNG,  
KaggleTestSnippet\_HCA\_1370\_f.41v.PNG,  
KaggleTestSnippet\_HCA\_1370\_f.17v.PNG



Ink blots or smudges

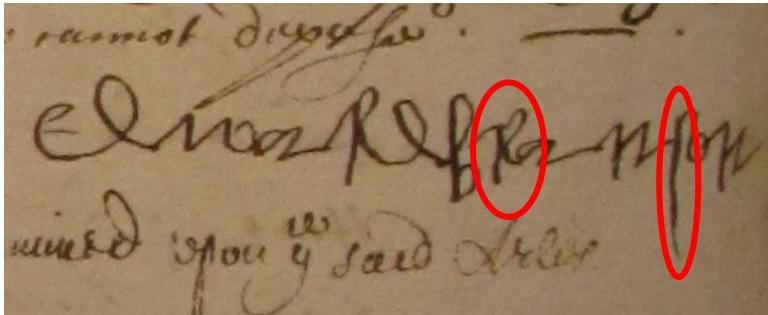
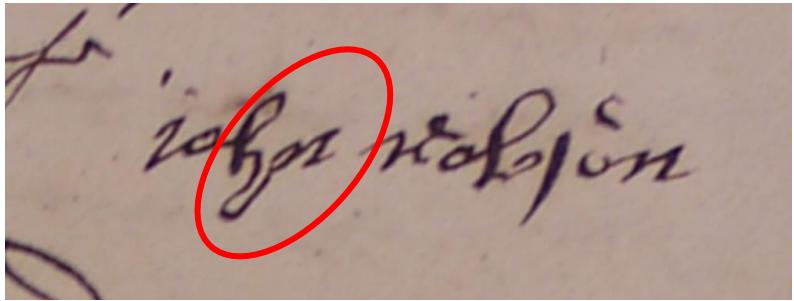


Stylistic feature or smudge?

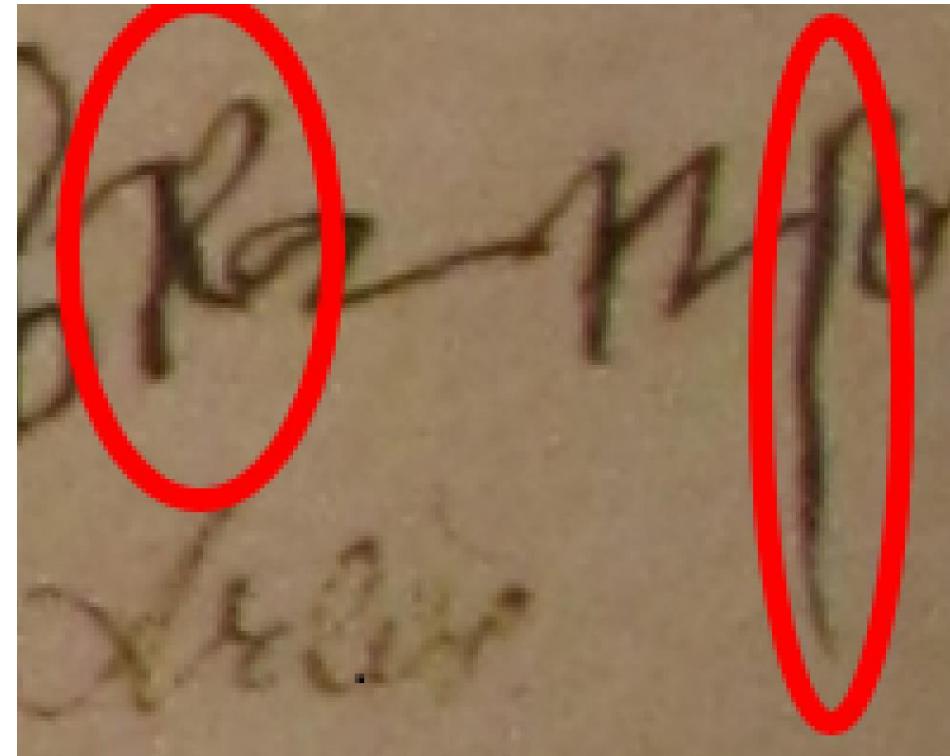
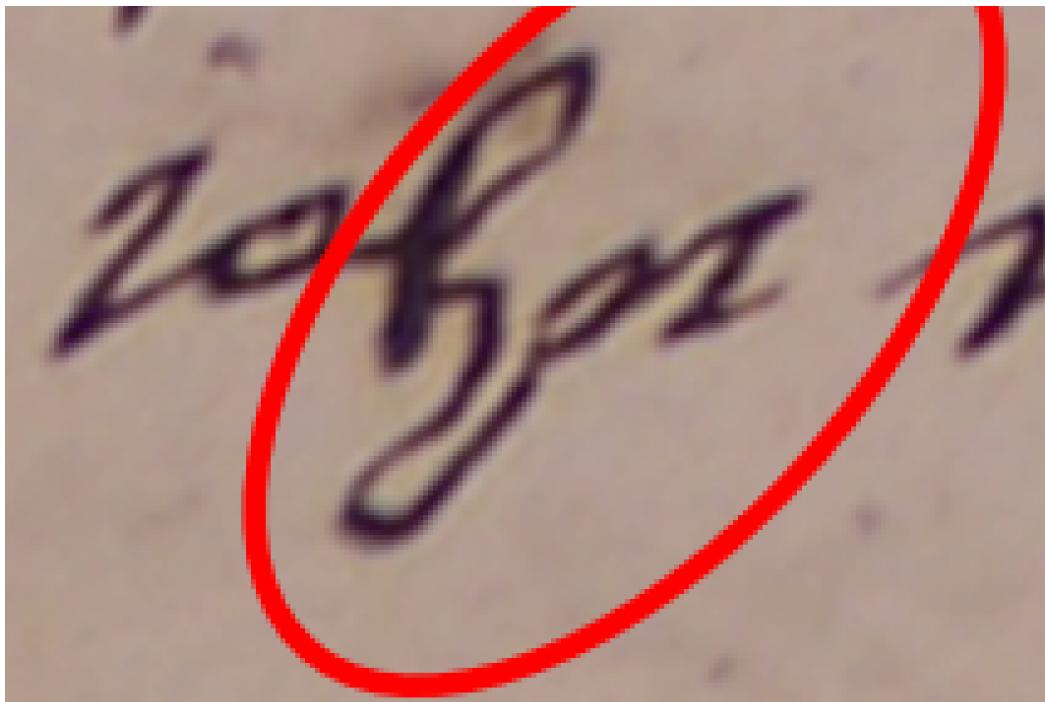


Deletion

## Detection of “shake” in straight and curved lines



HYPOTHESIS: Shaky lines may be a sign of poor signature execution (and by inference, poor handwriting execution) suggesting either lower level of literacy than smooth executed lines, or the effect of illness or age

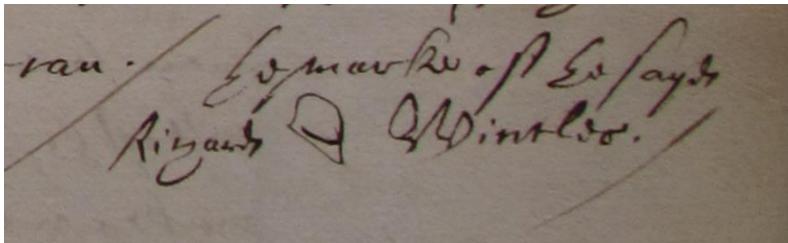


Shaky straight lines and/or loops

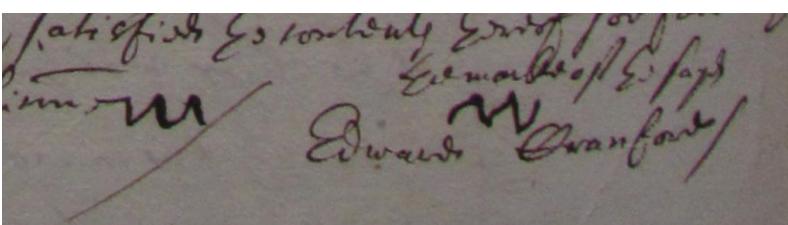
Clockwise, from top LH: KaggleTestSnippet\_HCA\_1371\_f.435v.PNG,  
KaggleTestSnippet\_HCA\_1368\_f.483v.PNG,  
KaggleTestSnippet\_HCA\_1368\_f.483v\_PIXELS.PNG,  
KaggleTestSnippet\_HCA\_1371\_f.435v.PNG\_PIXELS.PNG

# Porters handling coals, whale oil, ginger & corn

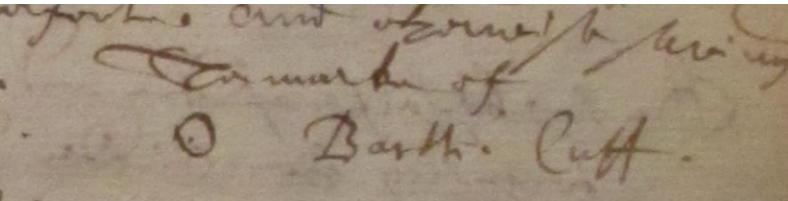
Richard Wincles, thirty-three year old porter, of the parish of Stepney, Middlesex, Dec. 15, 1656; employed as a labourer with fifteen other men to unload coals from the *Imployment* moored near Execution Dock, Wapping, into lighters for fixed rate of 12 s per man ([HCA 13/70 f.554r](#))



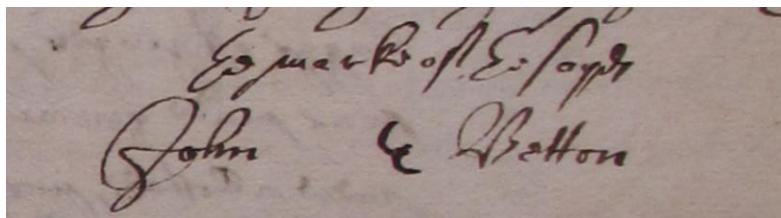
Edward Cranford, forty-four year old coale heaver or porter, of the parish of Stepney, Middlesex, Dec. 15, 1656; employed as a labourer with fifteen other men to unload coals from the *Imployment* moored near Execution Dock, Wapping, into lighters for fixed rate of 12 s per man ([HCA 13/70 f.555v](#))



Bartholomew Cuff, sixty year old porter of the Stillyard, of the parish of Allhallows the Greate, London, May 15, 1658; assisted in the landing of whale oil from lighters at the Stillyard Key and loading them away into a warehouse ([HCA 13/70 f.555v](#))



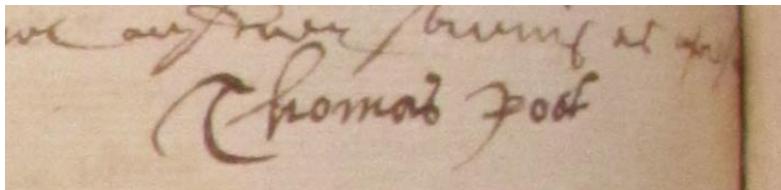
John Betton, fifty-four year old citizen and white baker of London, of the parish of Saint Buttolph Algate, London, Jul. 31, 1655; self-described as a porter employed by the Commissioners for Prize Goods to deliver ginger from a warehouse at Ralphes Key ([HCA 13/70 f.449r](#))



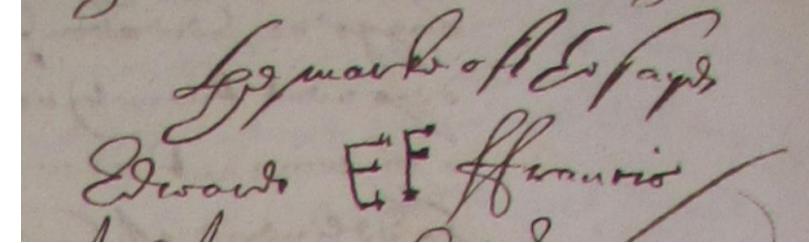
Edward Sherwin, fifty-six year old cittizen and leatherseller, of the parish of Little Allhallowes, London, Jul. 31, 1655; self-described as a porter employed by the Commissioners for Prize Goods to deliver ginger from a warehouse at Ralphes Key ([HCA 13/70 f.449v](#))



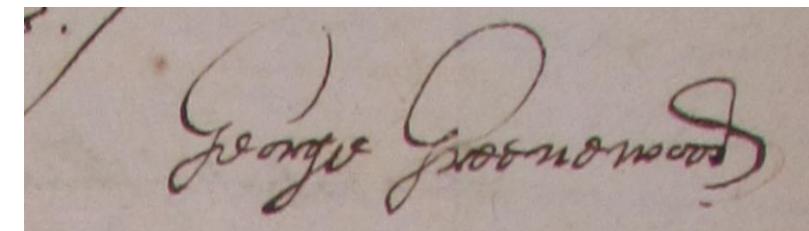
Thomas Roots, twenty-nine year old porter, of the parish of Greate Allhallowes, London, May 15, 1658; assisted in the landing of whale oil from lighters at the Stillyard Key, as one of the Stillyard porters, and loading them away into a warehouse ([HCA 13/72 f.330v](#))



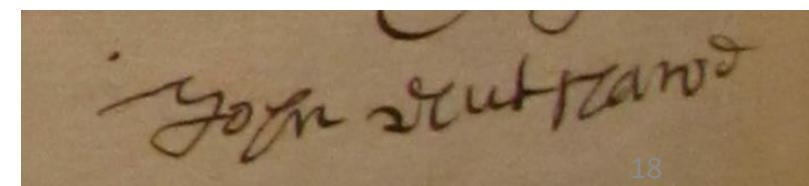
Edward ffrancis, citizen and merchant taylor of London, of the parish of Saint Olave in Southwarke, Jul. 31, 1655; self-described as a porter employed by the Commissioners for Prize Goods to deliver ginger from a warehouse at Ralphes Key ([HCA 13/70 f.450v](#))



George Greenwood, thirty year old citizen and vintner of London, of the parish of Saint Buttolph Bishopsgate, London, Jul. 31, 1655; self-described as a porter employed by the Commissioners for Prize Goods to deliver ginger from a warehouse at Ralphes Key ([HCA 13/70 f.454r](#))

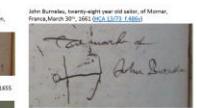
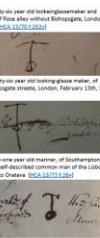
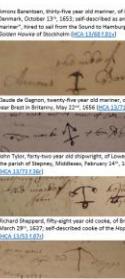


John Nutshall, fifty-five year old corne porter, of the parish of Saint Saviours Southwarke, Nov. 19, 1653; employed with a barber chyrurgeon/corne meter, an additional corne-meter, and other labourers to unlade a cargo of what in the *ffortune* of Stettin, moored against Limehouse; eight years of experience as a corne porter ([HCA 13/70 f.352v](#))



# SOLM-2018 IIIF anchors manifest in Mirador viewer

## Anchors



44

HCA Depositions: Anchors

The Mirador viewer interface displays a large central image of a historical manuscript page featuring several anchor markings. Below this main image, five smaller thumbnail images are shown, each corresponding to one of the anchor drawings. The thumbnails are labeled with the names of the individuals whose anchors they represent: Richard Shepperd, Andrew Beake (2), Andrew Beake (1), John Tylor, and John Burnelau. The 'Andrew Beake (1)' thumbnail is highlighted with a blue border, indicating it is the active or selected item. The interface includes standard Mirador navigation controls such as arrows for page navigation, zoom controls, and a search bar.

Richard Shepperd

Andrew Beake (2)

Andrew Beake (1)

John Tylor

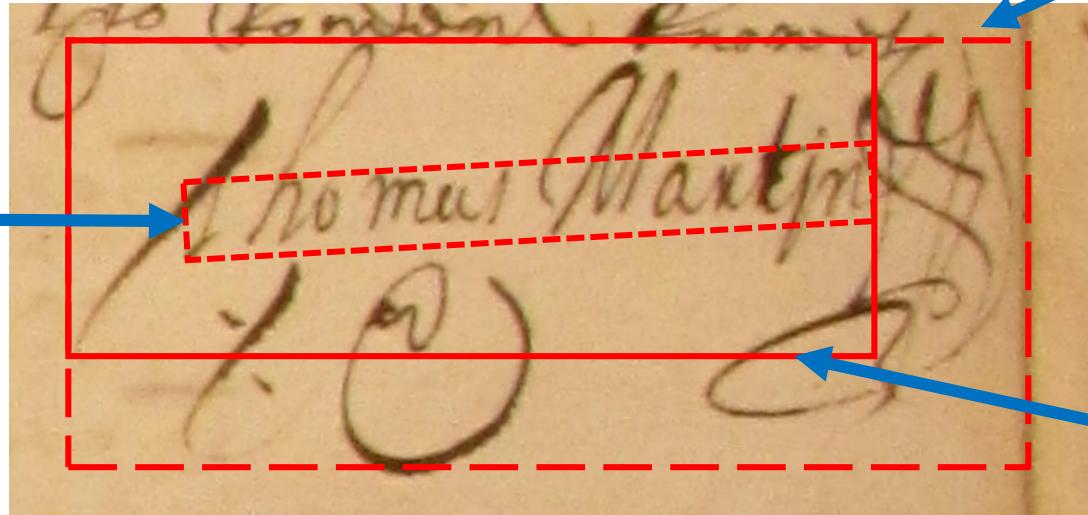
John Burnelau

# Boundary boxes marking the visual geometry of a signature

Inside boundary box,  
excluding uppers and  
downers

Outside boundary  
box, including  
flourish

Middle boundary  
box, including all  
letters, but excluding  
flourish



## Statistics

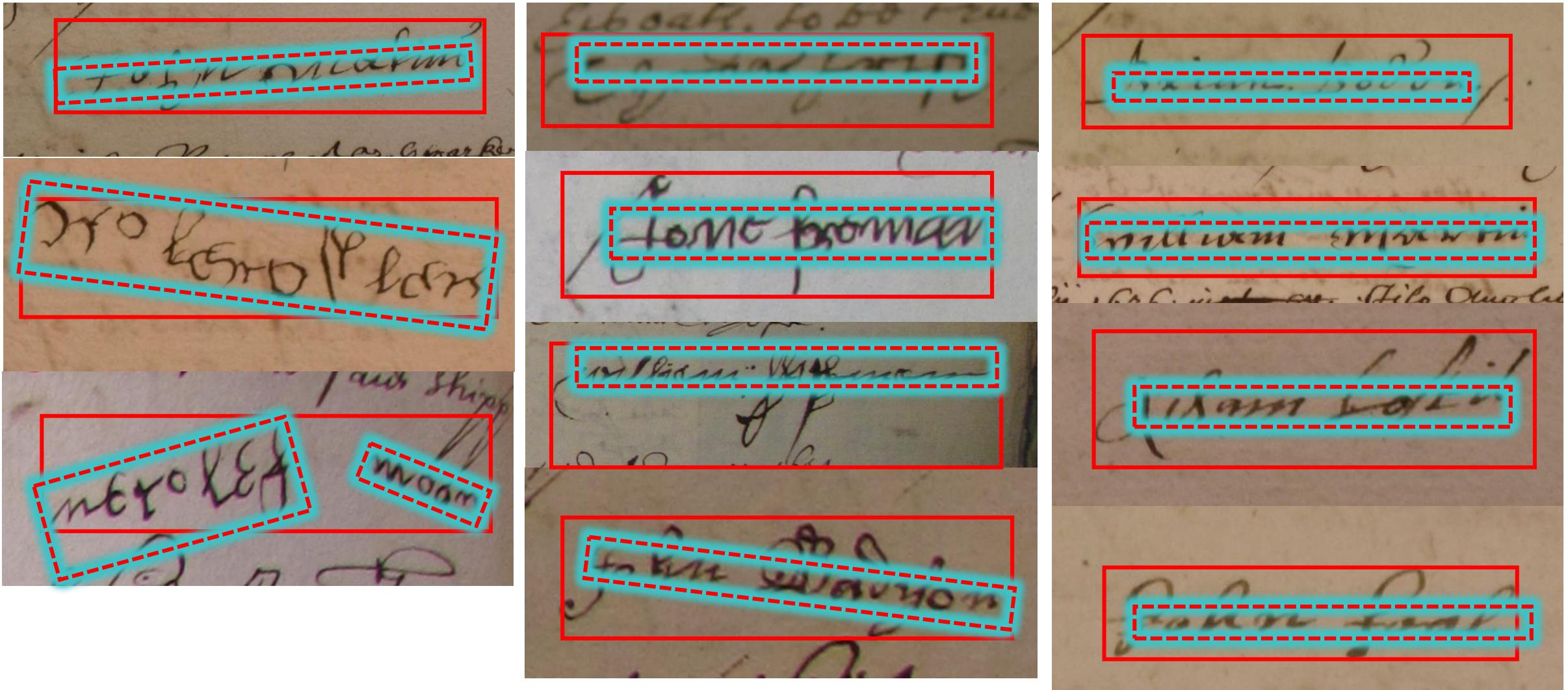
Inside boundary box: 9.0 x 1.1

Middle boundary box: 9.75 x 4.25

Outside boundary box: 12.75 x 5.75

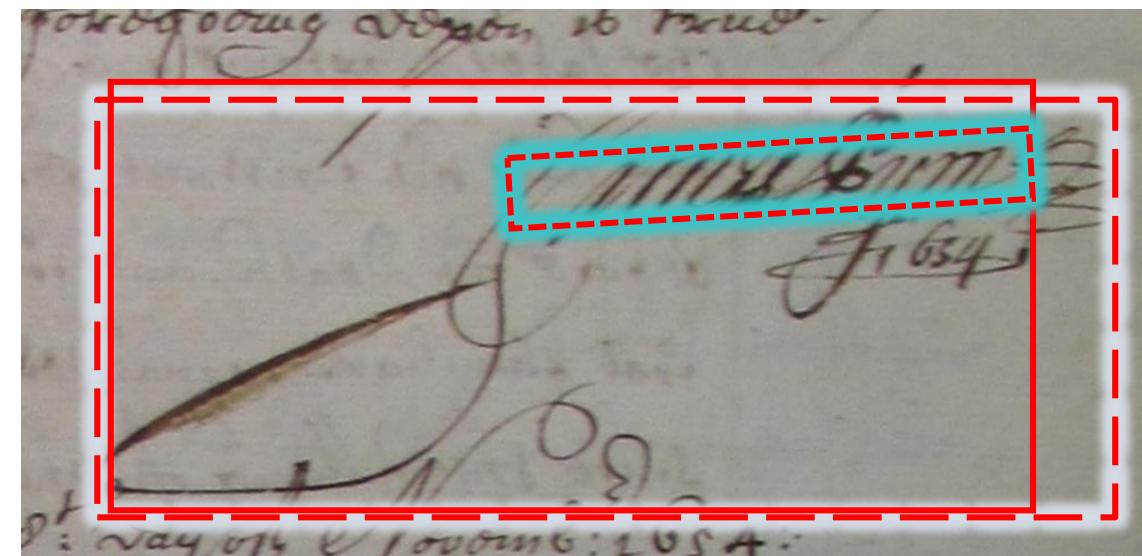
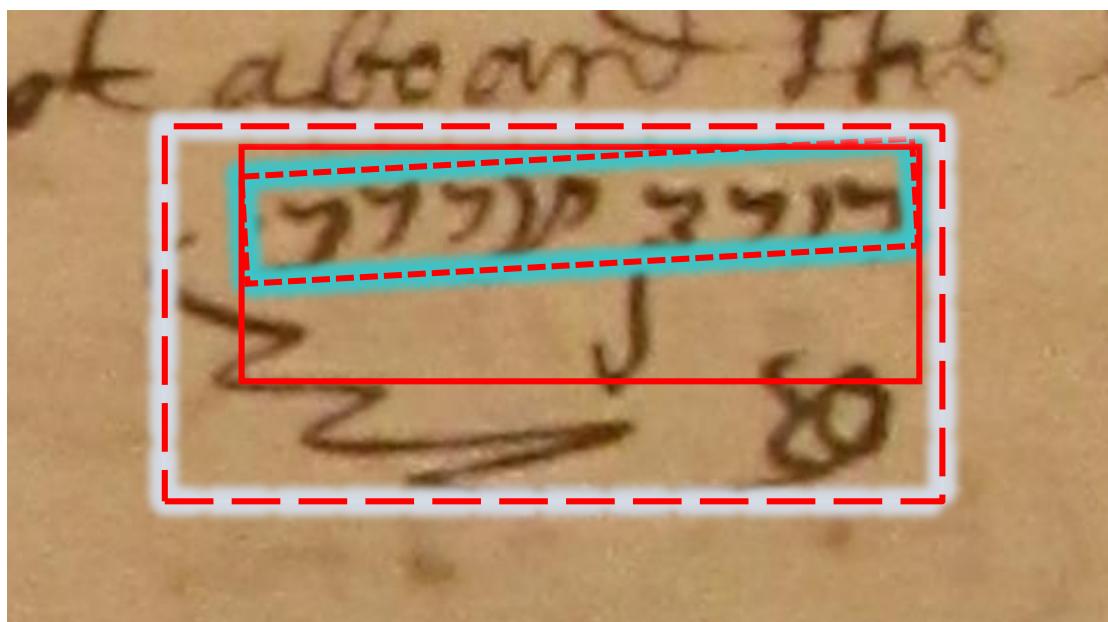
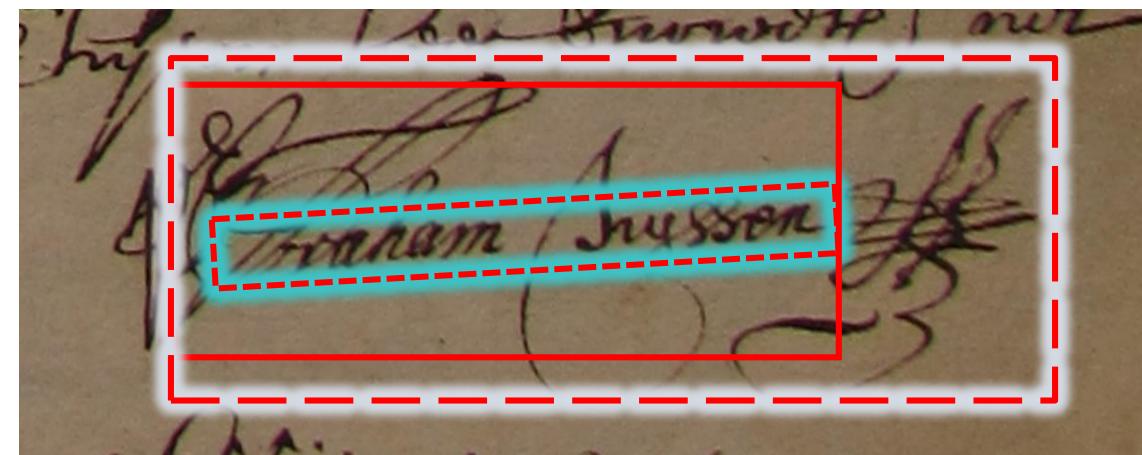
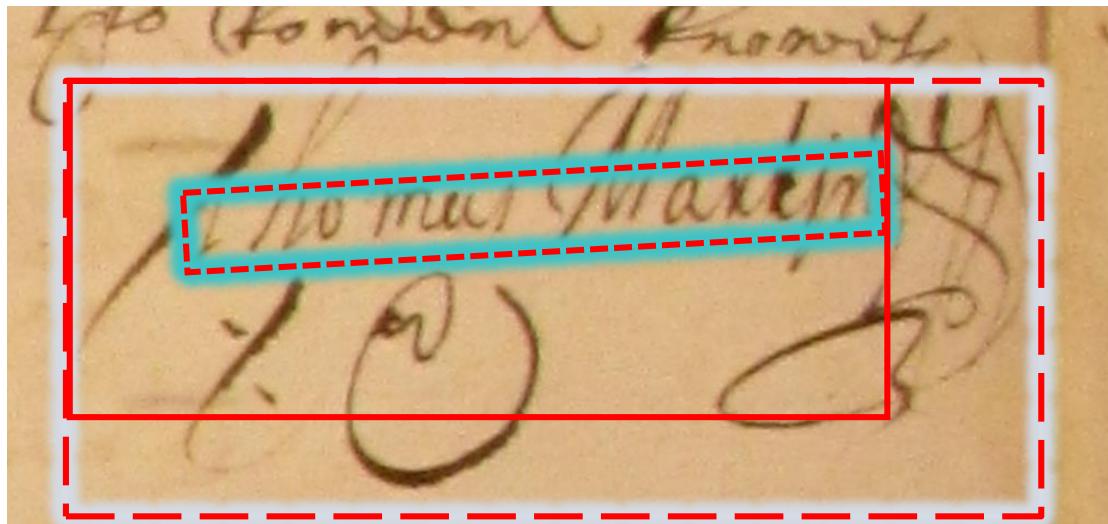
Rotation from horizontal: ca. 340 degrees

# Simple signatures, no flourishes



Source: Down from top LH side: KaggleTestSnippet\_HCA\_1353\_f.24v.PNG, KaggleTestSnippet\_HCA\_1353\_f.188r.PNG;  
Down from top Middle: KaggleTestSnippet\_HCA\_1353\_f.66r.PNG; KaggleTestSnippet\_HCA\_1370\_f.193r\_One.PNG,  
KaggleTestSnippet\_HCA\_1370\_f.203r.PNG, KaggleTestSnippet\_HCA\_1370\_f.218r.PNG  
Down from top RH SIDE: KaggleTestSnippet\_HCA\_1353\_f.28v.PNG, KaggleTestSnippet\_HCA\_1353\_f.29v\_One.PNG,  
KaggleTestSnippet\_HCA\_1353\_f.35r.PNG, KaggleTestSnippet\_HCA\_1353\_f.36v.PNG

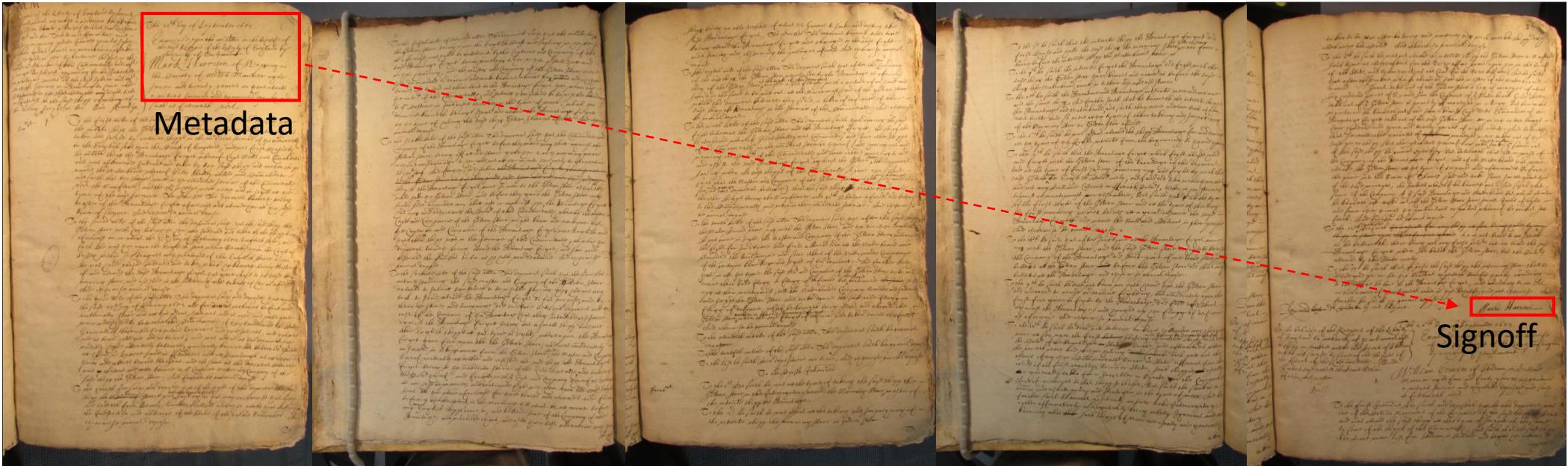
# Visual geometries of flourishes – C17th Irish, Dutch, English & Moroccan merchants



Source: Clockwise from top LH side: KaggleTestSnippet\_HCA\_1368\_f.34v.PNG,  
KaggleTestSnippet\_HCA\_1370\_f.366r.PNG, KaggleTestSnippet\_HCA\_1370\_f.134r.PNG,  
KaggleTestSnippet\_HCA\_1368\_f.58r.PNG

# Legal deposition

Deposition of Mark Harrison; mariner and master; resident in Wapping, Middlesex; age 27;  
Dated September 21<sup>st</sup> 1659 (TNA, HCA 13/68, ff. 1r-3r)



# Machine based recognition of metadata

The 21<sup>st</sup> Day of September 1693

Examined upon the 21<sup>st</sup> day of September 1693 before me by  
John Hobart Esq; a Notary Publick for the County of  
Norfolk and the City of Norwich, and by  
Mark Harrison of Wapping in  
the County of Middle Marches aged  
fifty and twenty years or thereabouts  
a mariner and seaman and seafarer  
and a native of Colchester and  
of the County of Essex.

On the first of the said month the day and year before  
the aforesaid John Hobart Esq; and Mark Harrison  
had taken and sailed on a certain shippe in the said County of Essex  
in the English seas upon the Coast of England, and had left Deptford  
the aforesaid shippe the Starreplough bound for Colchester  
and were of thence followed and taken by two other shipps and another  
named the Starreplough whereof Peter Hetherell master was commanded  
and forth that two shippes went in company to the port of Colchester  
with the Starreplough and the other shippes were never seen in sight as to  
time of the said shippes departing from Colchester to the port of  
Colchester. And upon the said day  
the said shippe of the said John Hobart Esq; and the said shippe the  
Starreplough with the Starreplough was followed and taken by the shippes  
abovesayd on or about the 25<sup>th</sup> day of February 1693 English year  
and said shippe was upon the English seas when it触礁 upon the Coast of  
England within the purview and jurisdiction of the English Admiralty  
and said shippe followed and taken up to the port of Colchester  
and about the said Starreplough said shippe was sent to the port of  
Colchester and was in the holding and taking of the aforesaid  
John Hobart Esq; and the said shippe.

On the third day of the said 1693 the day and year before  
the said 25<sup>th</sup> day of February 1693 also for divers reasons he first and  
subtlye chose and set by John Hobart Esq; and another man named  
Peter Hetherell to be his mate of the Starreplough the shippe  
General of Hartlepool bound for Colchester and so he  
arrived unto Colchester and so he arrived; and was sent to the port of Colchester and  
probably upon the 25<sup>th</sup> day of February 1693 he became the mate of the  
shippe General of Hartlepool and at London as aforesaid  
he was sent to the port of Colchester and the said John Hobart Esq;  
was appointed and made Captain of the Colchester Master and Company of  
the said shippe the Starreplough. And forthwith he named himself  
John Hobart Esq; and the said John Hobart Esq;

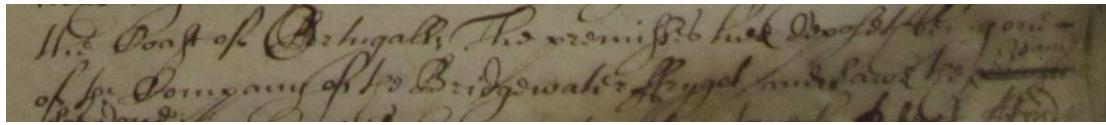
On the fourth day of the said 1693 and divers reasons he  
was the said John Hobart Esq; and so he became the captain of the said shippe  
and so he came to Colchester to be the captain and so he was sent to  
the port of Colchester and so he became the captain of the said shippe. And  
so he named himself John Hobart Esq;

The 21<sup>st</sup> Day of September 1689

Examined upon the affaile on the behalfs of  
the sayd Negroes of the Liberty of En gland by  
Mark T. Garrison of Newbury in  
the County of Middlesex aged  
seven and twenty years or there abouts  
sworn from and examined before me and  
signed as followeth. ver

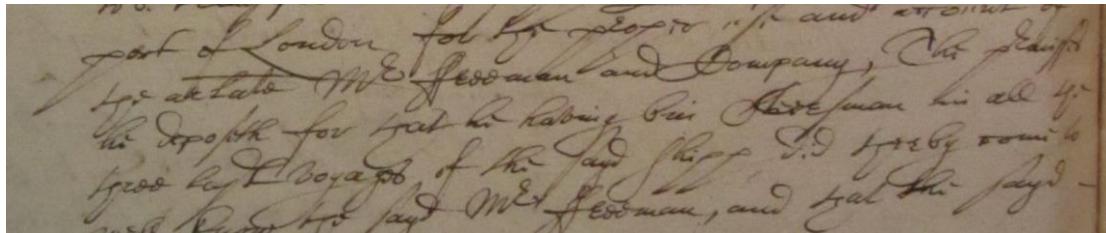
# Can we use key word spotting to excavate raw metadata?

## LANGUAGE DENOTING OCCUPATION



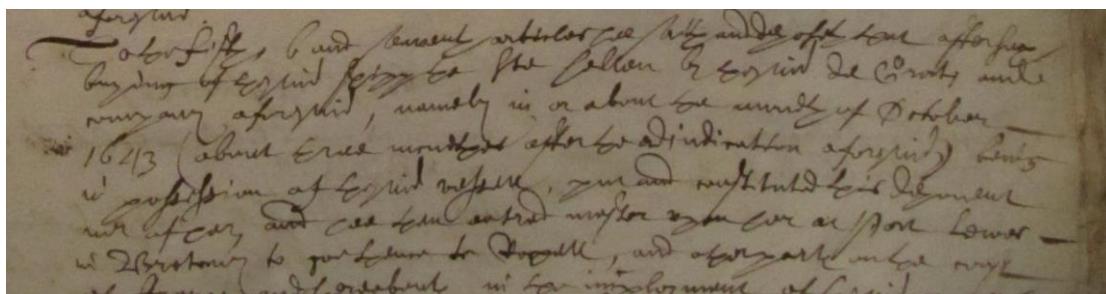
W<sup>t</sup> Roast of Porthgall, The premissee hee deposeth  
of the Companye of the Bridgewater ffrygott, and sawe her  
in the same shipp in the said port.

"The premisses hee deposeth being one of the company of  
the Bridgewater ffrygott, and sawe the same soe done" [HCA 13/72  
f.90r] [CONCLUSION: One of the company]



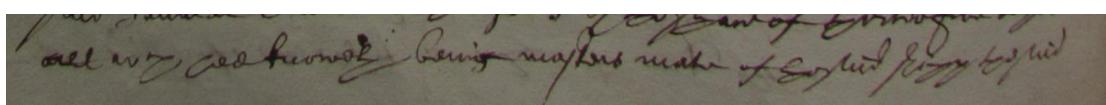
part of London for the seare of her and comon  
she ar late Mr. Godman and Company, the shipp  
the deppoth for that he had his Steersman in all ye  
free last boyards of the sayd shipp S. D. byt of comon  
to the sayd Mr. Godman, and that the sayd

"The premisses he deposeth for that he the deponent was not onely  
for the voyage arlate wherein she was stranded, but in two former  
voyages stiersman of the sayd ship" [HCA 13/72 f.90v] [CONCLUSION:  
Steersman]



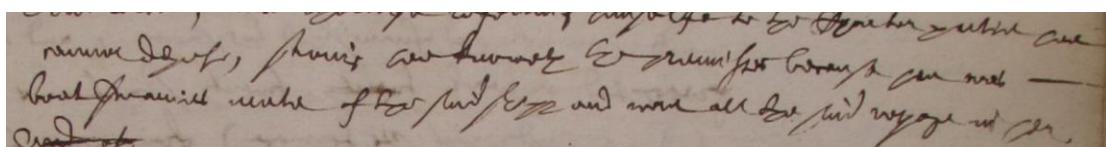
Yours.  
To oblyf, and present witness as he ande off late after  
buying of Esq<sup>r</sup> J<sup>r</sup> to the Bellin or Esq<sup>r</sup> de Grotte and  
many other apysit, namely in or about the moneth of October  
1643 (about three moneths after the adiudication aforesaid) being  
in possession at Esq<sup>r</sup> repre, and wrought bid him self  
out of say and see from entred master of the shipp at Port Lewes  
in Bretany to go to her to Foggall, and afterward in the my  
lone to Looe in the moneth of Nov<sup>r</sup>.

"after such buying of the said shipp the *Santa Hellen* by the said da  
[?Groots] and company aforesaid, namely in or about the moneth of  
October 1643 (about three monethes after the adiudication aforesaid)  
being in possession of the said vessel, put and constituted this  
deponent master of her, and hee then entred master upon her at Port  
Lewes in Bretany" [HCA 13/72 f.95r] [CONCLUSION: Master]



all my ded knowell being masters mate of sayd shipp by mid

"all which hee knoweth being masters mate of the said shipp the said  
voyage" [HCA 13/70 f.669v] [CONCLUSION: Master's mate]



name deput, shipp ar knowne to greate ande  
boatmane mate of sayd shipp and made all sayd shipp and her  
creed etc

"hee knoweth the premisses because hee was boatswaines mate of the  
said shipp and went all the said voyage in her" [HCA 13/70 f.671r]  
[CONCLUSION: Boatswain's mate]

Can we refine raw machine generated metadata using a combination of NPL, controlled vocabularies, and programmable decision rules?

# LANGUAGE DENOTING OCCUPATION

"The premisses hee deposeth being one of the company of the *Bridgewater ffrygott*, and sawe the same soe done"  
[HCA 13/72 f.90r] [CONCLUSION: One of the company]

“The premisses he deposeth for that he the deponent was not onely for the voyage arlate wherein she was stranded, but in two former voyages stiersman of the sayd ship” [HCA 13/72 f.90v] [CONCLUSION: Steersman]

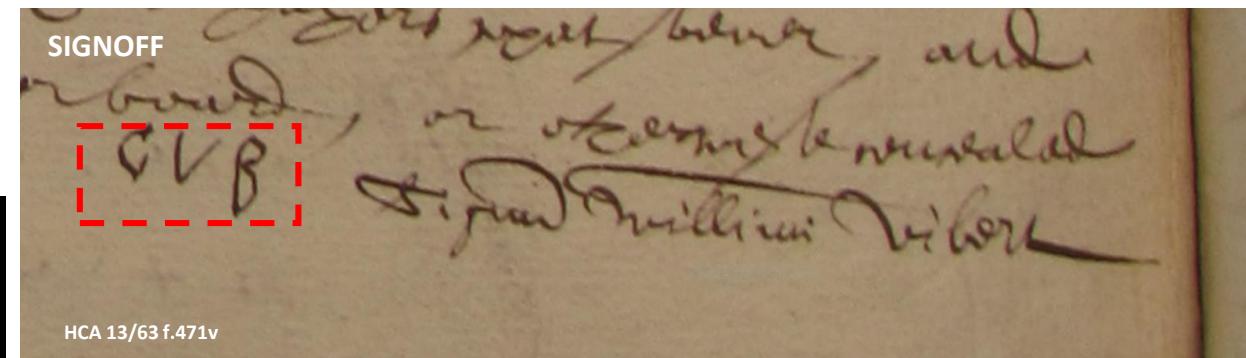
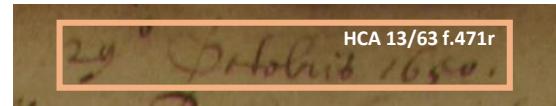
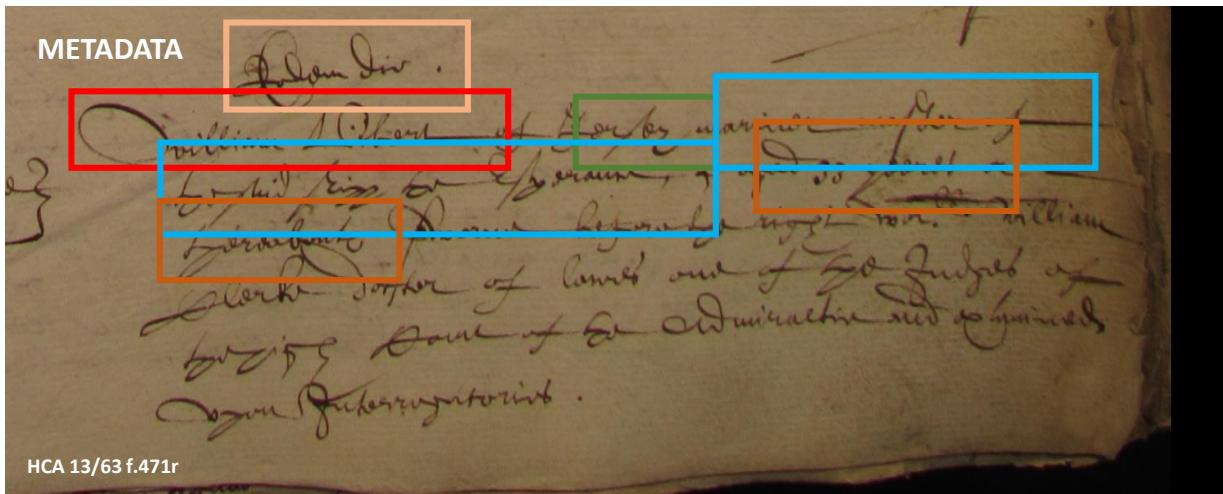
"after such buying of the said shipp the *Santa Hellen* by the said da [?Groots] and company aforesaid, namely in or about the moneth of October 1643 (about three monethes after the adjudication aforesaid) being in possession of the said vessel, put and constituted this deponent master of her, and hee then entred master upon her at Port Lewes in Bretany" [HCA 13/72 f.95r] [CONCLUSION: Master]

"all which hee knoweth being masters mate of the said shipp the said voyage" [HCA 13/70 f.669v] [CONCLUSION: Master's mate]

"the premisses because hee was boatswaines mate of the said shipp and went all the said voyage in her" [HCA 13/70 f.67r] [CONCLUSION: Boatswain's mate]

KaggleTestSnippet_HCA_1370_f.546r.PNG	HCA 13/70	Signature	Mariner; Boatswain
KaggleTestSnippet_HCA_1370_f.571v.PNG	HCA 13/70	Signature	Mariner; Boatswain
KaggleTestSnippet_HCA_1370_f.596v_One.PNG	HCA 13/70	Signature	Mariner; Boatswain
KaggleTestSnippet_HCA_1370_f.636r.PNG	HCA 13/70	Signature	Mariner; Principal boatswain
KaggleTestSnippet_HCA_1370_f.671v.PNG	HCA 13/70	Marker	Mariner; Boatswain's mate
KaggleTestSnippet_HCA_1368_f.631v.PNG	HCA 13/68	Signature	Mariner; Boatswain
KaggleTestSnippet_HCA_1371_f.27r.PNG	HCA 13/71	Initials	Mariner; Boatswain
KaggleTestSnippet_HCA_1371_f.27v_One.PNG	HCA 13/71	Initials	Mariner; Boatswain
KaggleTestSnippet_HCA_1371_f.27v_Two.PNG	HCA 13/71	Initials	Mariner; Boatswain
KaggleTestSnippet_HCA_1368_f.640r.PNG	HCA 13/68	Signature	Mariner; Boatswain
KaggleTestSnippet_HCA_1368_f.657r.PNG - CREATE HCA 13/68	HCA 13/68	Signature	Mariner; Boatswain [of the Civil Society]
KaggleTestSnippet_HCA_1371_f.77v.PNG	HCA 13/71	Signature	Mariner; Boatswain
KaggleTestSnippet_HCA_1370_f.378r.PNG	HCA 13/70	Signature	Mariner; Boatswain
KaggleTestSnippet_HCA_1371_f.99r.PNG	HCA 13/71	Signature and	Mariner; Boatswain [of man of war]
KaggleTestSnippet_HCA_1370_f.484v.PNG	HCA 13/70	Signature	Mariner; Quartermaster; Boatswain
KaggleTestSnippet_HCA_1371_f.139v.PNG	HCA 13/71	Signature	Mariner; Boatswain
KaggleTestSnippet_HCA_1371_f.167r.PNG	HCA 13/71	Signature	Mariner; Boatswain [of the John and Mary]
KaggleTestSnippet_HCA_1371_f.279r.PNG	HCA 13/71	Signature	Mariner; Boatswain

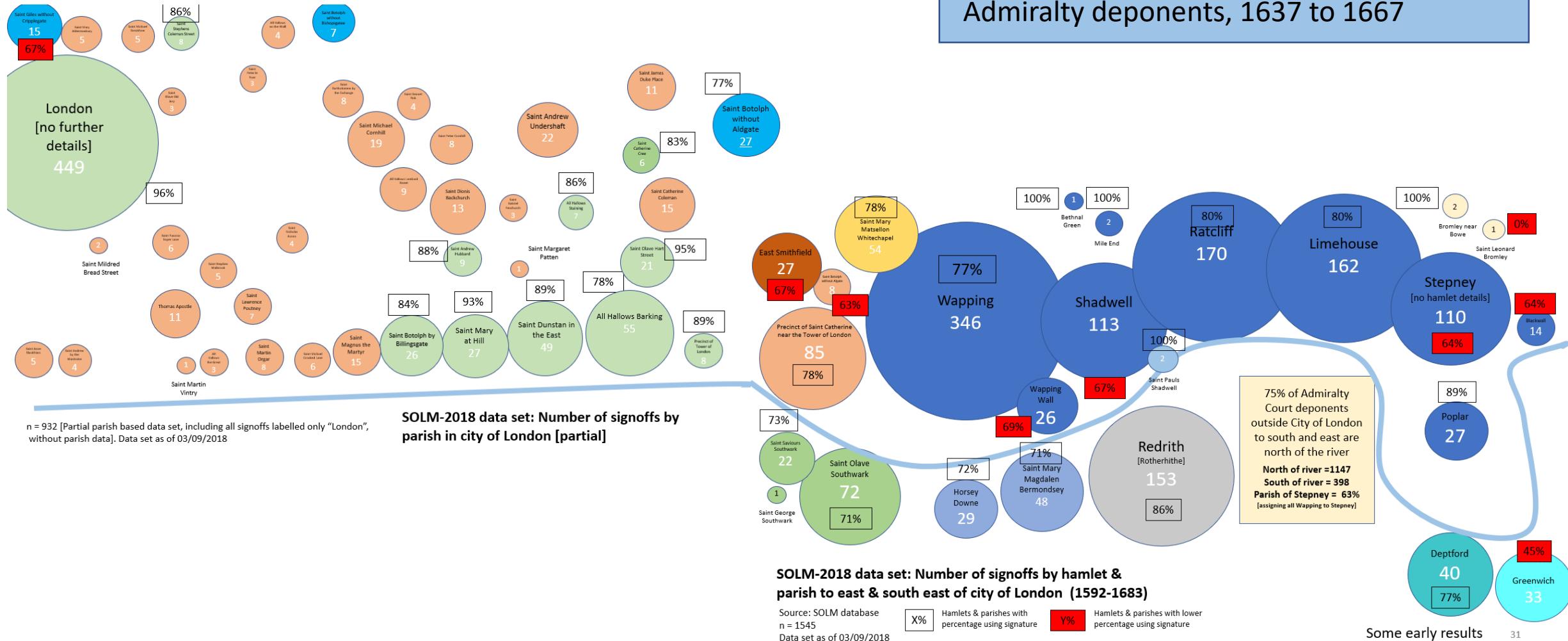
Imagine an ArchiveBot extracting metadata automatically  
from handwritten manuscripts and working with volunteers to finalise



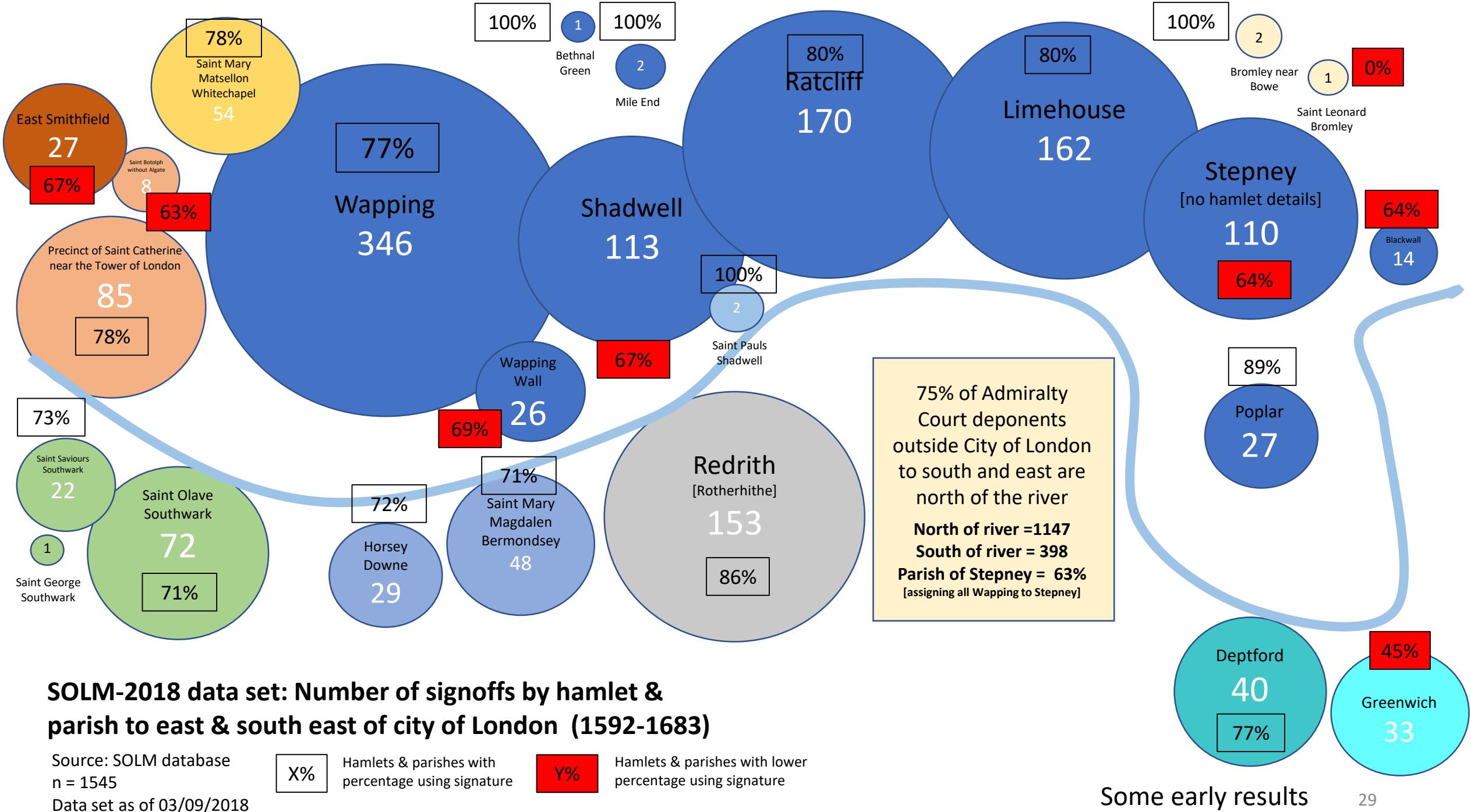
	Date
	Name
	Residence
	Occupation
	Age
	Signoff

## Some early results

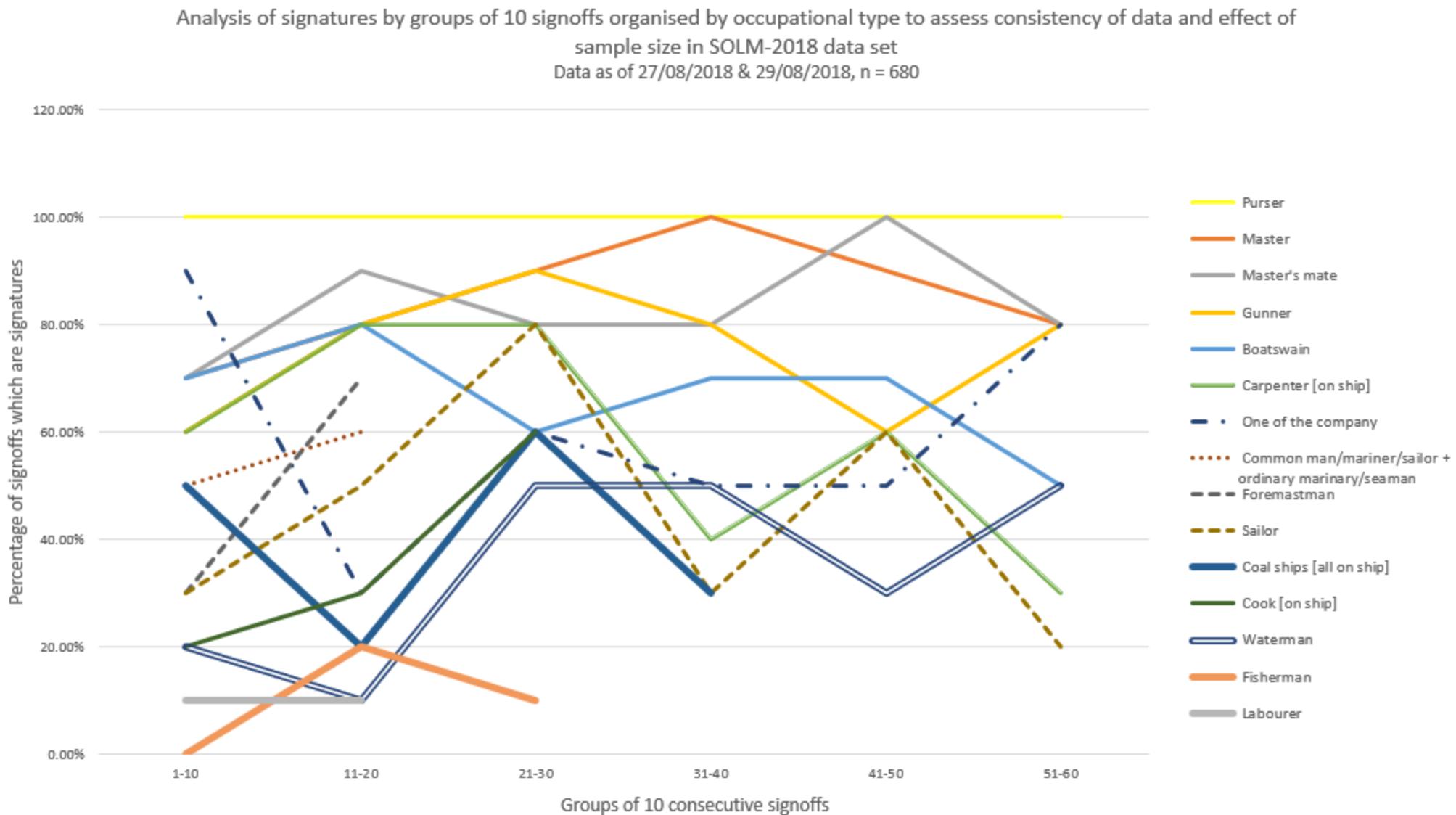
Early/mid-C17th London – a linear maritime city, as seen in the location of High Court of Admiralty deponents, 1637 to 1667



Source: SOLM-2018 database



# Some early results



# Contact details

**Colin Greenstreet  
Community organiser,  
Signs of Literacy**

**Dr Mark Hailwood  
Lecturer in History, 1400-1700,  
University of Bristol**

Email:

[colin.greenstreet@gmail.com](mailto:colin.greenstreet@gmail.com)  
[m.hailwood@bristol.ac.uk](mailto:m.hailwood@bristol.ac.uk)

Weblinks:

<http://signsofliteracy.org>  
<http://marinelives.org>  
<http://chronoscopic.org>

GitHub:

<https://github/Signsofliteracy/Signoff>

Twitter:

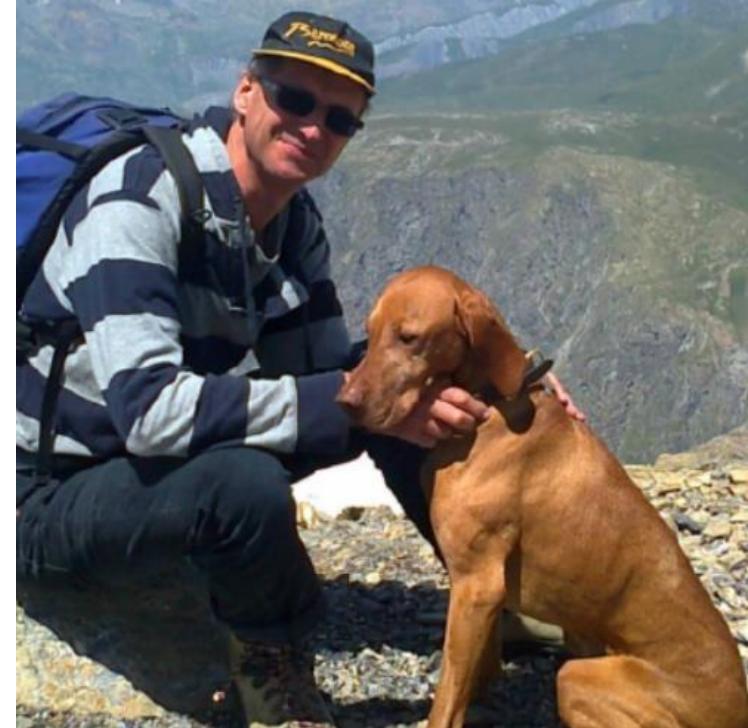
[Marinelivesorg](https://twitter.com/Marinelivesorg)

Working  
with:



The National Archives

kaggle Competitions



# **Discussion**

# For discussion at DHC 2018, Thursday, September 6<sup>th</sup>, 2018

	2018	2019	2020	2021	2022	2023
Q1		Historical literacy network meeting (1)				
Q2	IIIF Conference Staatsarchiv Amsterdam workshop DH Benelux	Historical literacy network meeting (2)				+ <div style="border: 1px solid black; padding: 10px;"><p>Machine learning enabled manuscript archivists</p><p>Image processing</p><p>Natural Language Programming</p><p>Key word spotting</p><p>Controlled vocabularies</p><p>Programmable decision rules</p><p>IIIF visibility</p></div>
Q3	Archives & AI Sheffield DH Congress	Kaggle competition Huntington library visit				
Q4	Kaggle visit Stanford University visit Transkribus Vienna users conference					
	<b>SOLM-2018</b>	<b>SOLM-2019</b>	<b>SOLM-2020</b>	<b>SOLM-2021</b>	<b>SOLM-2022</b>	<b>SOLM-2023</b>

# **Backup material**

Table 6.6 *Ranking of trades by illiteracy in rural England, 1580–1700*

Trade	No. sampled	No. mark	% mark
Scrivener	21	0	0
Apothecary	15	0	0
Vintner	13	0	0
Ironmonger	10	0	0
Mercer	32	1	3
Draper	46	2	4
Grocer	60	3	5
Haberdasher	14	1	7
Merchant	122	12	10
Dyer	14	2	14
Clothier	86	18	21
Goldsmith	12	3	25
Baker	48	13	27
Innkeeper	36	10	28
Glazier	29	8	28
Saddler	17	5	29
Chandler	16	5	31
Barber	30	10	33
Tanner	101	37	37
Brewer	42	16	38
Maltster	34	13	38
Woolcomber	40	18	45
Mariner	28	13	46
Weaver	524	257	49
Wheelwright	34	17	50
Fuller	30	15	50
Victualler	14	7	50

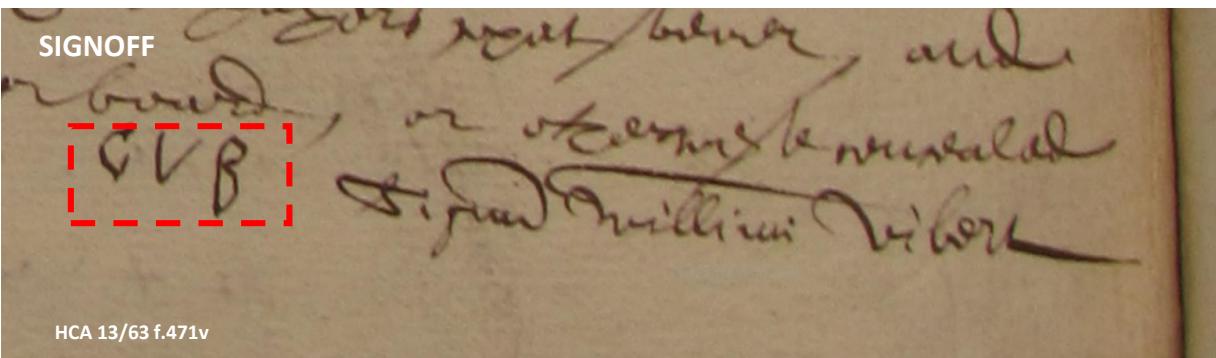
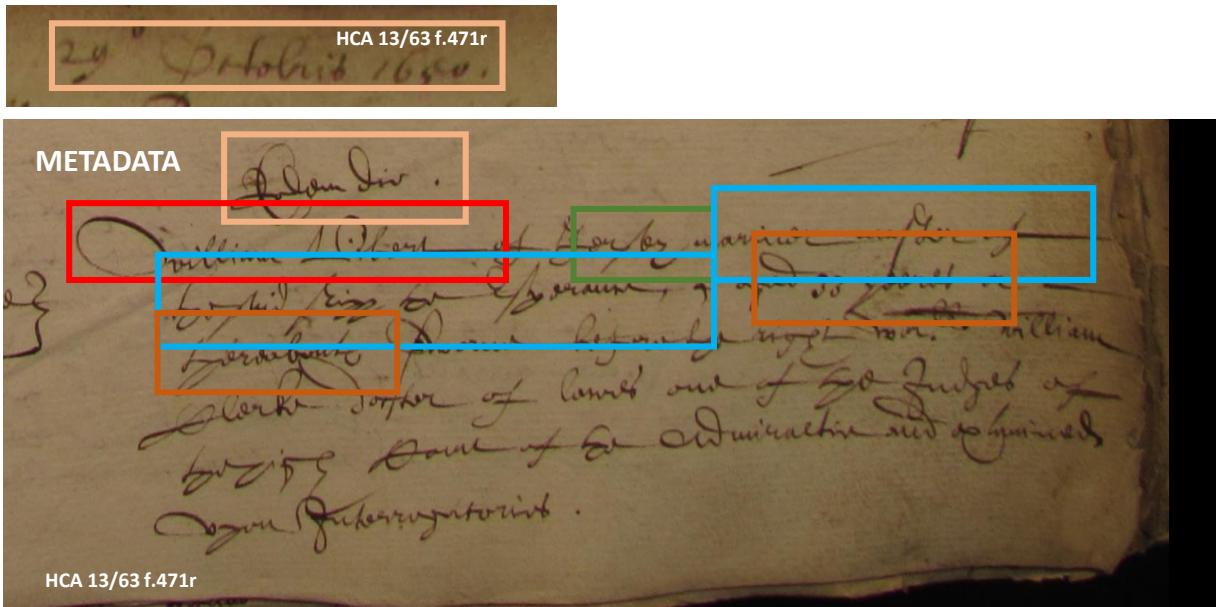
Table 6.6 cont.

Trade	No. sampled	No. mark	% mark
Tailor	286	145	51
Cordwainer	135	70	52
Turner	11	6	55
Blacksmith	137	77	56
Sailor	40	23	58
Worsted-dresser	33	19	58
Joiner	26	15	58
Wright	27	16	59
Butcher	157	94	60
Miller	44	27	61
Carpenter	201	124	62
Sherman	13	8	62
Glover	72	45	63
Gardener	22	14	66
Shoemaker	57	37	65
Cooper	44	30	68
Bricklayer	45	33	73
Collier	15	11	73
Currier	11	8	73
Mason	57	44	77
Cutler	10	8	80
Ropemaker	10	8	80
Hellier	21	17	81
Fisherman	16	13	81
Shepherd	11	9	82
Thatcher	43	39	91
Miner	25	24	96
Slater	11	11	100

Table 7.3 *Illiteracy of tradesmen and craftsmen in the dioceses of Durham, Exeter, Norwich and London, 1560–1730*

Decade	Durham/ Northumberland				Devon/Cornwall				Norfolk/Suffolk				Essex/Herts.				London/Middlesex			
	No. sampled	No. mark	% mark		No. sampled	No. mark	% mark		No. sampled	No. mark	% mark		No. sampled	No. mark	% mark	No. sampled	No. mark	% mark		
1560–9	92	77	84																	
1570–9	153	117	76	100	70	70														
1580–9	53	42	79						98	60	61	60	41	68	134	55	41			
1590–9	79	45	57						161	89	55					101	43	43		
1600–9	132	73	55						151	72	48	68	34	50	65	13	20			
1610–19	116	65	56	55	23	42	126	55	44	73	16	22	172	44	26					
1620–9	102	51	50							93	34	37	212	63	30					
1630–9							140	69	49	113	42	37	154	50	32					
1640–9							90	47	52											
1660–9				177	79	45	176	58	33											
1670–9				314	151	48	149	52	35					171	37	22				
1680–9				215	89	41	174	77	44					222	58	26				
1690–9							125	38	30					167	21	13				
1720–9							104	35	34					170	14	8				

# Imagine an ArchiveBot extracting metadata automatically from handwritten manuscripts and working with volunteers to finalise



	Date
	Name
	Residence

	Occupation
	Age
	Signoff



William Vibert (Guillaume Vibert) was a thirty-three year old mariner from the isle of Jersey, who was master of the ship the *Esperansa*.

He signed his deposition in the English High Court of Admiralty on October 29<sup>th</sup> 1650, using three not two initials.

The choice of “G”, “V” and “B” for initials suggests he thought of himself and pronounced his name as “Guillaume Vee Bert” (though he could not write a full signature)

## Challenge for a Bot

- (1) Find the date of the deposition (it is at the top of the same manuscript page above the preceding deposition)
- (2) Deal with the interlineation in the metadata
- (3) On the next manuscript page, recognise “GVB” are initials not a mark. The “G” is quite tough (easy to mistake for “C”). Yes, they are recognisable letters, but the three don’t match the two names “William Vibert”. Need to know William – Guillaume in French, and that mariners in Jersey may be born in France and French speaking.

We need visual metadata, which can be machine processed

**Table 1.2a EXPANDED: HCA 13/53 [f.1r-340v] - Signoff frequency per manuscript page, data from 1637**

	1 r	2 v	3 r	4 v	5 r	6 v	7 r	8 v	9 r	10 v	Subtotal									
1-10	1	2	3	1	0	1	0	2	2		16									
11-20	2	2	1		1	1	1	1	1	1	13									
21-30	2	1		1	3	1	1	1	2	1	16									
31-40		1	1	1		1	1	2	1	1	13									
41-50		1		1			1	1	1		6									
51-60			1		1	2	1		2	2	11									
61-70	2			1	1		2	1	1	1	18									
71-80	1		2	1		1	2	1	1	1	19									
81-90	2	1	1	1	1	2	1	2	4	1	23									
91-100	1	2			1	2	2	3	3	1	26									
101-110	2	1	2	2	1	1	2	2	1	2	23									
111-120	1			1		1	1	2	1	2	16									
121-130			1			2		1	1	2	12									
131-140	2	3	2	1	1	2		1	1	1	1	23								
141-150	1	1	2	1	2	2	1	2	1	1	22									
151-160			1			2	1	1	2	1	18									
161-170		1		2	2	1		1	2	1	17									
171-180	1	2			2	1	1	1		1	11									
181-190				2				3	1	2	1	21								
191-200		1			1	1	1	1	1	1	17									
201-210	2	2			1	1	2	4	3	1	24									
211-220	1			2	1		1	1	4	1	25									
221-230	2	2	1		3		1	1	1	2	25									
231-240	1	1		1		2	1	1	3		15									
241-250	2						2	1	1	2	1	15								
251-260	2		2	2	1		1	1	1		1	15								
261-270		1	1		1	1		1	1	1	2	11								
271-280	2			1		1	1	1		1	1	12								
281-290	1		1		2	1	1	1	1	2		14								
291-300		1	1	1	1	2		1	1	1	1	15								
301-310	1			2			2	1	2	1		12								
311-320			1			1			1	2		6								
321-330				1			1			1	3	7								
331-340	1	2	2	2		2	1	1	1	1	1	18								
Total	31	30	24	27	23	17	23	30	19	31	39	31	33	30	42	33	16	29	16	555

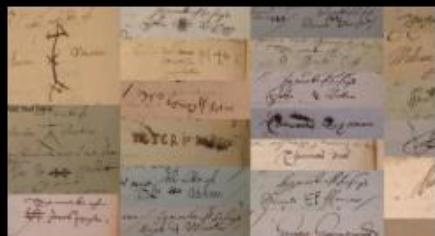
# Project portfolio

<http://www.chronoscopic.org>

## MarineLives



## Signs of Literacy



## EM Textiles, Garments & Dyestuffs Glossary



## Maphackathon



## SOLM-2018



## EM Maritime & Mercantile Gazetteer



# Thanks to our one hundred and ninety-six contributors since 2012

We would like to recognise and thank our friends and supporters who have contributed to the MarineLives project, to the Maphackathon event, to the Early Modern Textiles, Garments and Dyestuffs glossary, and to the Signs of Literacy initiative, whether as volunteer transcribers, annotators, commentators, glossary contributors, software developers, advisors, interviewees, workshop participants, workshop and conference speakers, or as PhD Forum participants (in alphabetical order)

Dr Aquiles Alencar-Brayner	Thomas Davies	Viveka Hansen	M. L. Logue	Benjamin Redding	Dr Samantha Thompson
Dr Kimberly Alexander	Jonathan Dent	Elaine Harrington	Jelle van Lottum	Irmila Regulova	Roger Towner
Dr Aaron Allen	Melvyn Dresner	Dr Liam Haydon	Ismail Malik	Ethan Reynolds	Dr Alexis Truax
Elizabeth Ames	Dr Stuart Dunn	Phillipa Hellawell	Grace Mallon	aniel Richards	Dr William Tullett
Dr Roberta Anderson	Professor Kai Eckert	Laura Herbert	Paula Marmor	ariet Richardson	Martin Turner
Vicky Annand	Bob Egan	Dr Helmer Helmers	Simon Marsh	ndrew Richens	Oliver Turner
Dr Carolyn Arena	Eglantine	Dave Henderson	Dr Alan Marshall	n Mia Ridge	David Underdown
Deborah Ashby	Dr Charlene Eska	Dr Stefan Hessbrüggen-Walter	Dr Barbara McGillivray	n Dominique Ritzé	Rebecca Unsworth
Dr Gary Baker	Louise Falcini	Professor Tim Hitchcock	Dr Angela McShane	n Gavin Robinson	Dr Brodie Waddell
Gabor Bakos	Emilie-Jane Farrimond	Dr Philip Hnatkovich	Angela Middleton	en Robson	Rebecca Want
Iona Bartzi	Dr Karwah Fatah-Black	Rachel E. Holmes	John Miller	n Helen Rogers	Samuel Watson
Rachel Bates	Dr Amelia Fay	Jeffrey Hopper	Anne Mills	n Jill Rogers	Dominic Webb
Rowan Beentje	Dr Janet Few	Dr Jenni Hyde	Kate Morant	cel Hee Rømer	Ken Whittaker
Nga Bellis-Phan	Christian Folgar	Dr Tom Irvine	Matthias Müller-Prove	eline Romano	Jill Wilcox
Michael Bennett	Sara Fox	Steve Ives	Professor Steve Murdoch	n Margaret Schotte	Peter Williams
Amber Berkeley	Dr Anthony Firth	Alex Jackson	Dr Shavana Musa	n even Schrum	Royline Williams-Fontenelle
Dr Richard Blakemore	Erin French	Stefan Jäggi	Maurice Nicolson	nđrei Seleznov	Ad van der Zee
Lior Blum	Dr Ian Friel	Elin Jones	Stephanie Ostrich	ura Seymour	Dr Kathrin Zickermann
Ffion Boyd	Professor Cheryl Fury	Sue Jones	Frances Owen	Inn Sheridan	Dr Suze Zijlstra
Katie Broke	Steve Garnett	Menno Jonker	Gordon O'Sullivan	n Deborah Sherlock	Dr Cäcilia Zirn
Dr James Brown	Dr Perry Gauci	Dr Sam Kaislaniemi	Dr Katherine Parker	n Rainer Simon	and the ever helpful but
Dr Andy Burn	Marja Geesink	Ross Keel	Tim Parry-Williams	a Sjoberg	anonymous @_mapnut
Elio Calcagno	Jaap Geraerts	Dr Patricia Keller	David Pashley	ntr Sloup	
Joshua Calloway	Helen Good	William Kellett	Dr Michael Pearce	n Edmond Smith	
Michelle Maria Early Capistrán	Dr Jamie LH Goodall	Sara J Kerr	Dr Cathryn Pearce	Illy Smith	
Rachel Carter	Guy Grannum	Heather Knight	Elizabeth Pimblett	n Stephen Snelders	
Cynthia Chin	Colin Greenstreet	Iris Kramer	Dr Sophie Pittman	ofessor Humphrey Southall	
Nicola Clarke	Francesca Greenstreet	Dr Marcin Krygier	Professor Simone Paolo	n Craig Spence	
Freyalyn Close-Hainsworth	Justin Grierson	John Kuhn	Ponzetto	n Daniel Stewart-Roberts	
Dalton Coker	Tracey E Griffiths	Dr Craig Lambert	Mark Ponte	n Hannah Steyne	
Dr Giovanni Colavizza	Dr Adam Grimshaw	John Layt	Dr Peter Pridal	n Stoll	
Dr Justin Colson	Karen Gunnell	Deborah Leem	Dr Jo Pugh	ichael Stoner	
Sam Cottrell	Yerevag Hagopian	Dr Sjoerd Levelt	Seif Al Rashidi	n Chad Stolper	
Kevin Daniels	Dr Mark Hailwood	Amalia Skarlatou Levi	Dr Peter Rauxloh	eter Taylor	
Thierry Daunois	Finn Halligan	John Levin	Patrizia Rebullia	n Carl Thompson	
Dr John Davies		M. L. Logue			