Revisiting historical literacy

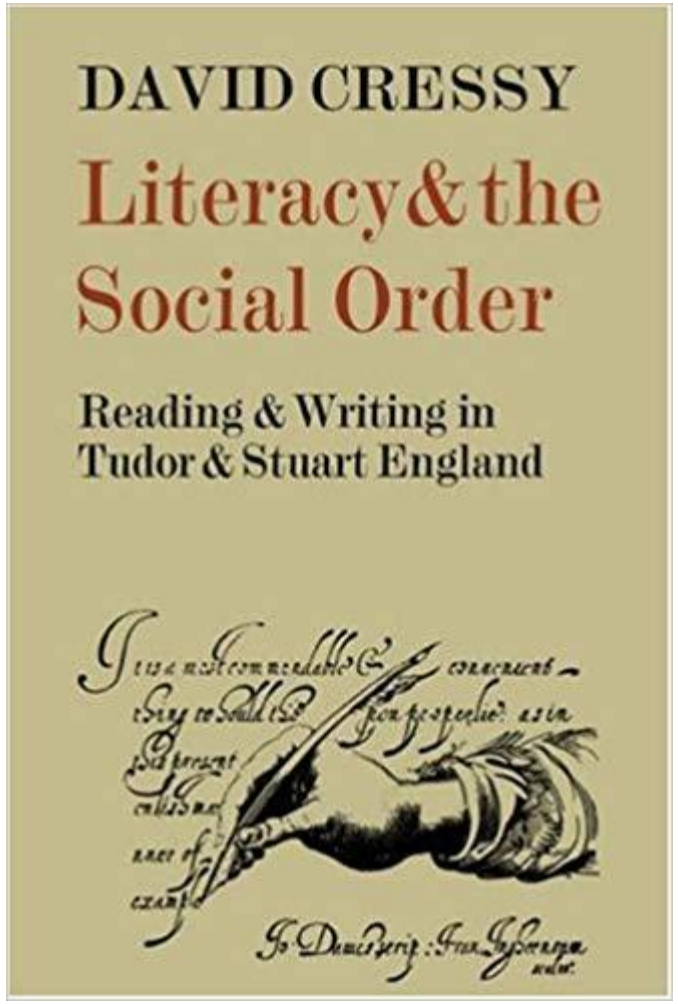
DHC 2018

Thursday, September 6th, 2018
University of Sheffield

Dr Mark Hailwood
Colin Greenstreet

Digital Humanities and Early Modern History

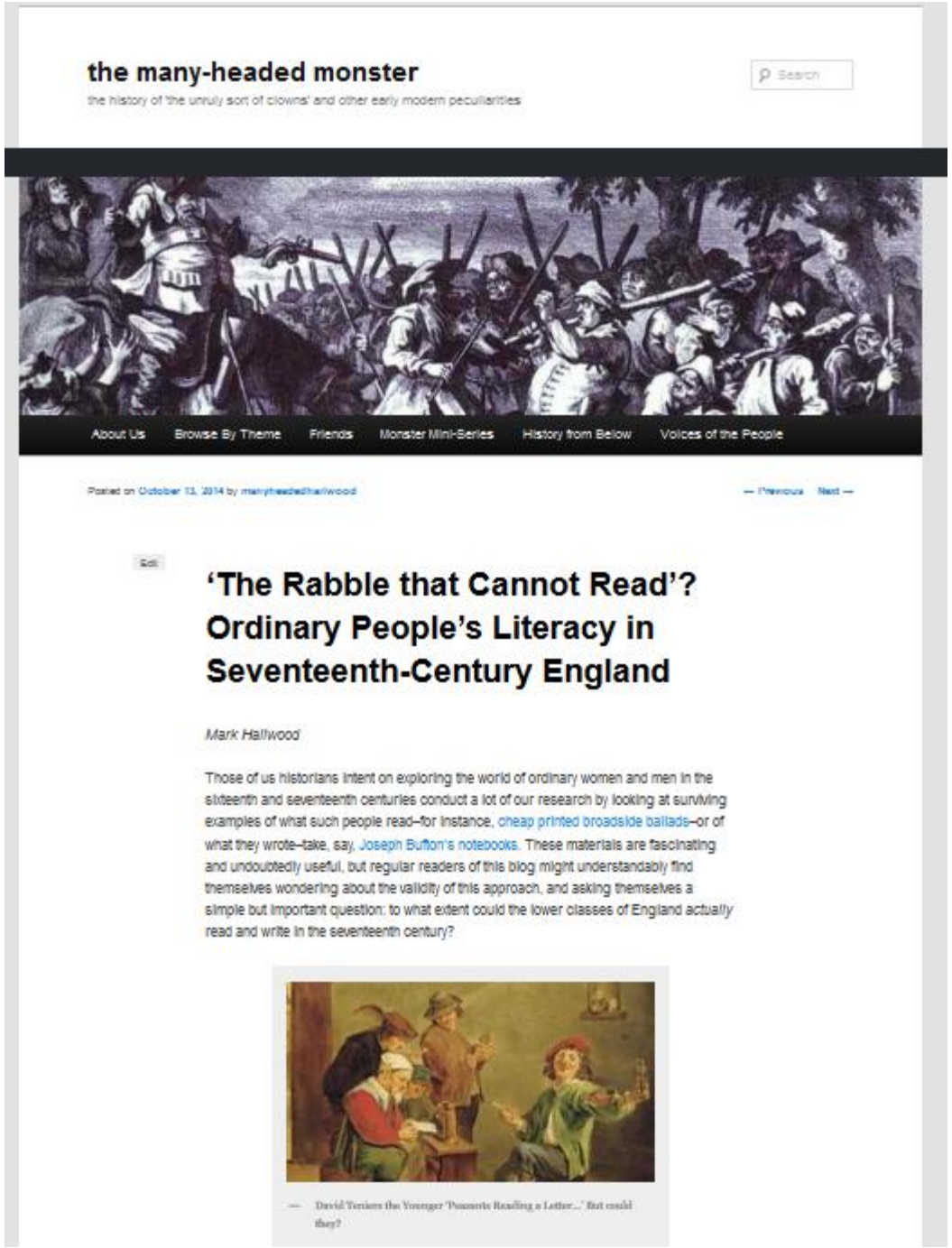
Who could read?



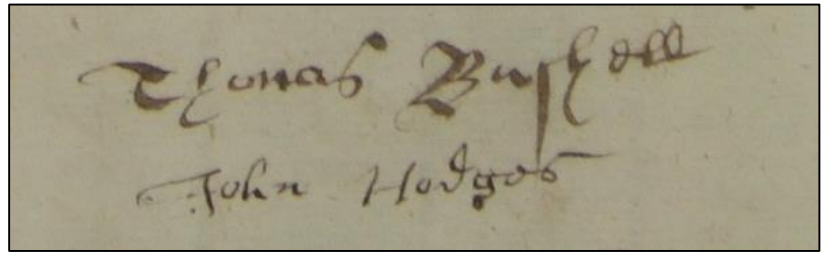
1980

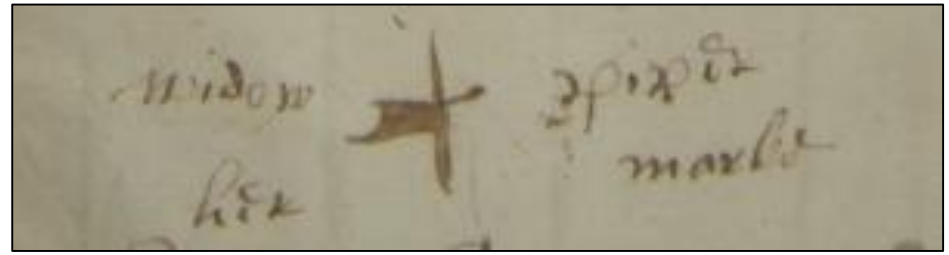
Table 1: Levels of illiteracy in early modern England (Cressy, 1980)

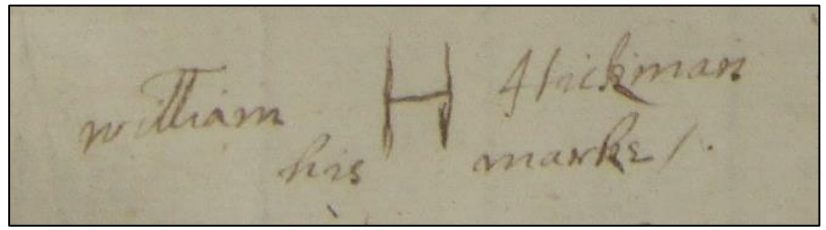
Date	Men	Women
1485–1507	90%	99%
1558	80%	95%
1642–1645	70%	90%
1714	55%	75%

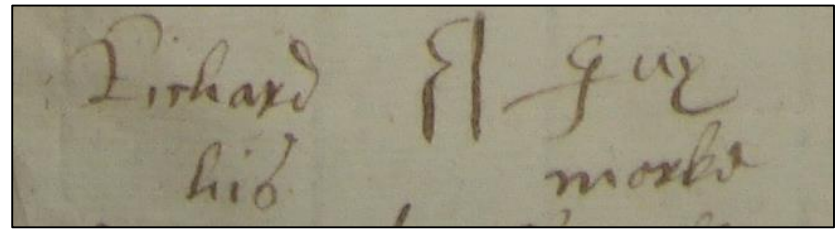


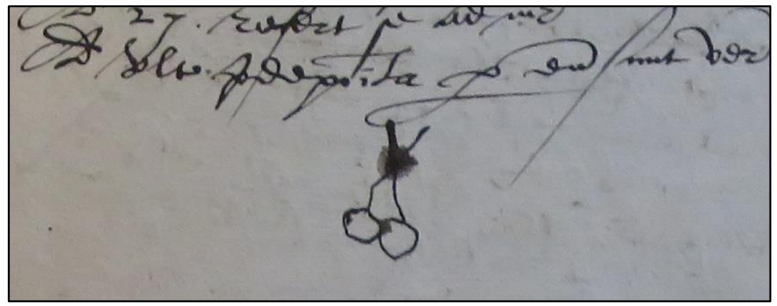
Source:
https://manyheadedmonster.wordpress.com/2014/10/13/the-rabble-that-cannot-read-ordinary-peoples-literacy-in-seventeenth-century-england/

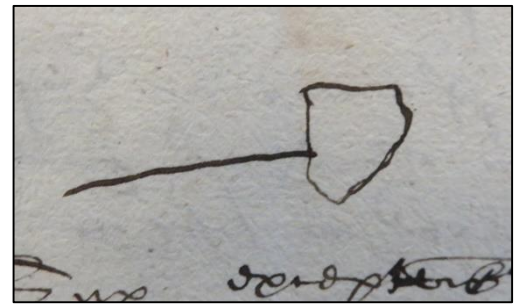


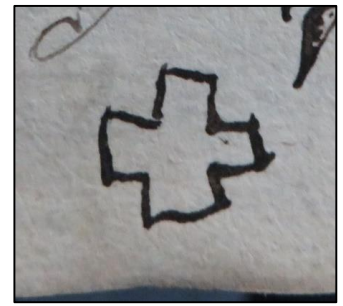


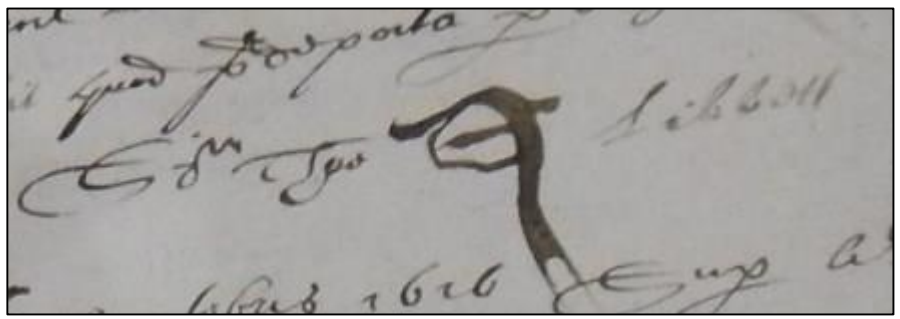


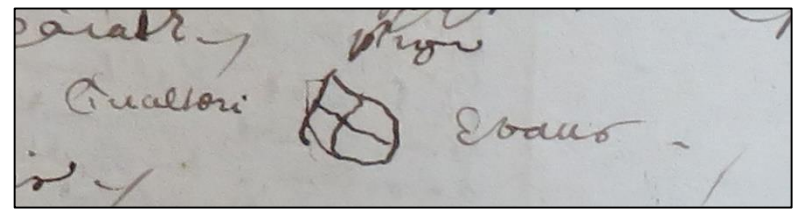


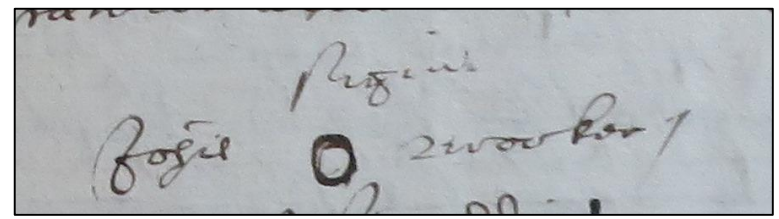












Journal of British Studies 54 (July 2015): 553–577. doi:10.1017/jbr.2015.61

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

Eleanor Hubbard

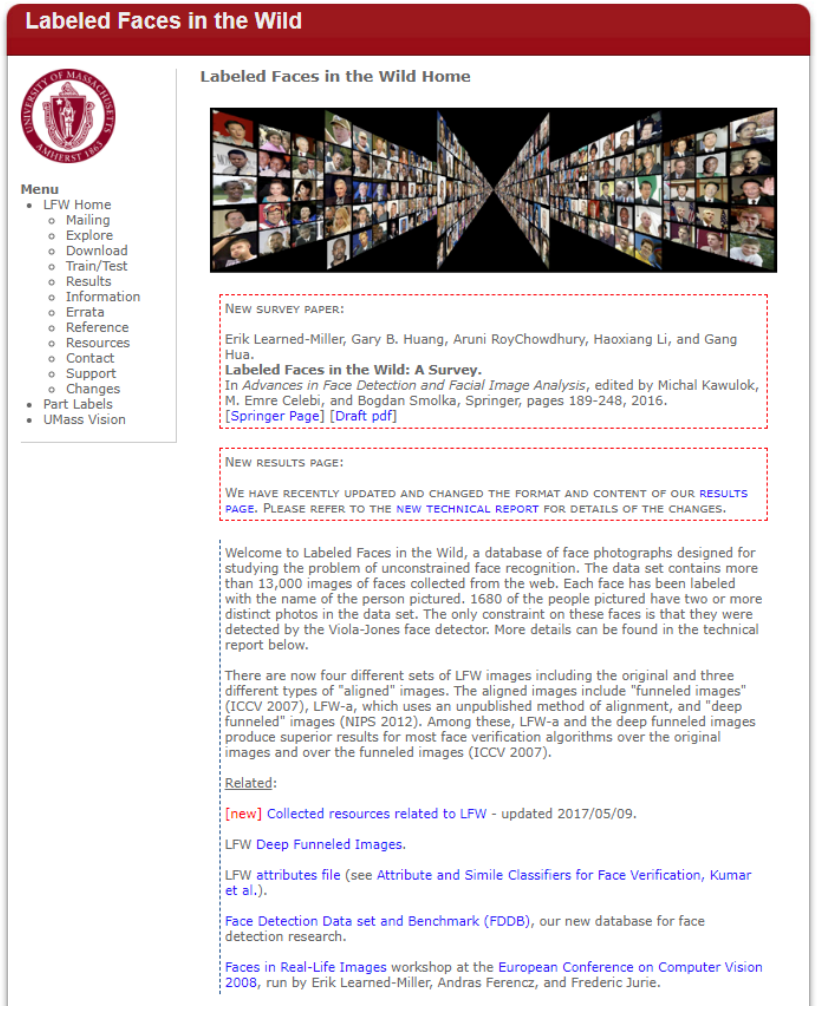
Abstract This article reopens the vexed question of how many women in early modern England could read by calling attention to the precise ways in which women marked, initialed, and signed legal depositions in late sixteenth- and early seventeenth-century London. It shows that initialing and signing were 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."[1] 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.[2] 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



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

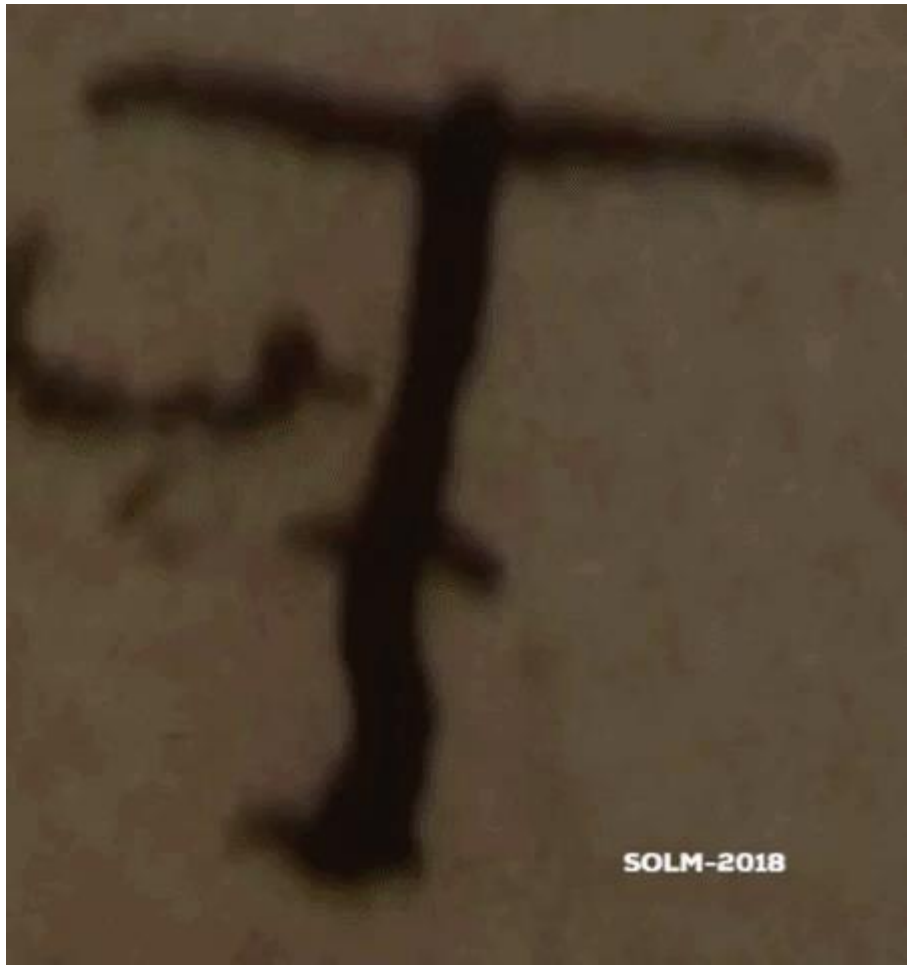
Erik Learned-Miller, Gary Huang, Aruni RoyChowdhury, Haoxiang Li, Gang Hua

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

Source: http://vis-www.cs.umass.edu/lfw/

Pattern seeking

How many different letters can you recognise?



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

Source: SOLM-2018 database

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

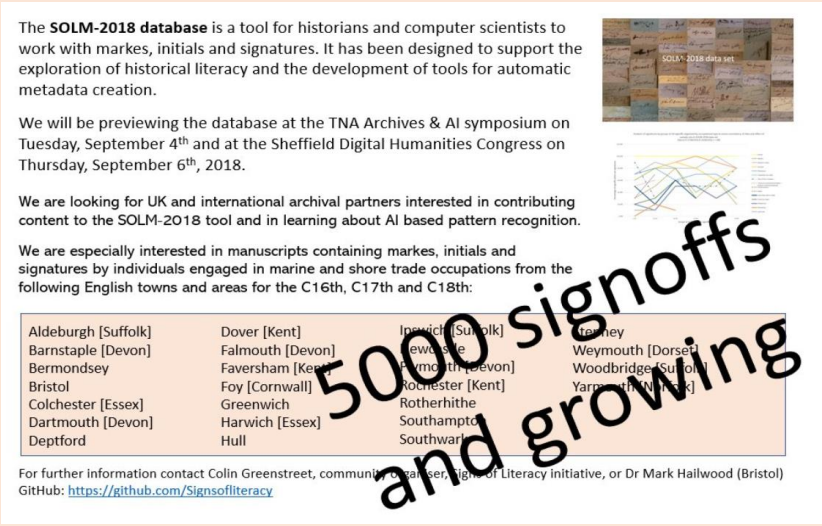
Source: SOLM-2018 database

Our vision is a SOLM-2023 database with 1 million markes, 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 markes, 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 4th and at the Sheffield Digital Humanities Congress on Thursday, September 6th, 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 markes, 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]	Walmey
Barnstaple [Devon]	Falmouth [Devon]	Lewes [Sussex]	Weymouth [Dorset]
Bermondsey	Faversham [Kent]	Plymouth [Devon]	Woodbridge [Suffolk]
Bristol	Foy [Cornwall]	Rochester [Kent]	Yarmouth [Norfolk]
Colchester [Essex]	Greenwich	Rotherhithe	
Dartmouth [Devon]	Harwich [Essex]	Southampton	
Deptford	Hull	Southwark	

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

50,000 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

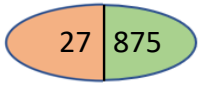
Cressy (1980) sample sizes in brown; SOLM-2018 (n=5,000) sample sizes in green
SOLM-2023 goal (n = 1 mill)

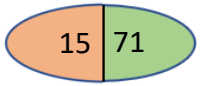
Table 6.7 *Ranking of London and Middlesex trades by illiteracy, 1580–1700*

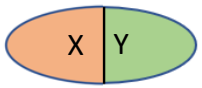
Trade	No. sampled	No. mark	% mark
Scrivener	46	0	0
Merchant	27	0	0
Vintner	23	0	0
Grocer	21	0	0
Saddler	10	0	0
Apothecary	9	0	0
Goldsmith	29	1	3
Stationer	18	1	6
Chandler	28	2	7
Barber	13	1	8
Ironmonger	13	1	8
Draper	34	4	12
Haberdasher	49	7	14
Dyer	13	2	15
Glazier	13	2	15
Leatherseller	12	2	17
Skinner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

Table 6.7 cont.

Trade	No. sampled	No. mark	% mark
Baker	19	5	26
Cook	11	3	27
Victualler	40	11	28
Sailor	17	5	29
Clothworker	30	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
Carpenter	40	16	40
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







Brown = Cressy (1980)
Green = SOLM-2018 (2018, half complete)

Source: David Cressy, Literacy and the Social Order (Cambridge, 1980), pp.134-135

Signs of Literacy Kaggle Research Competition

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:

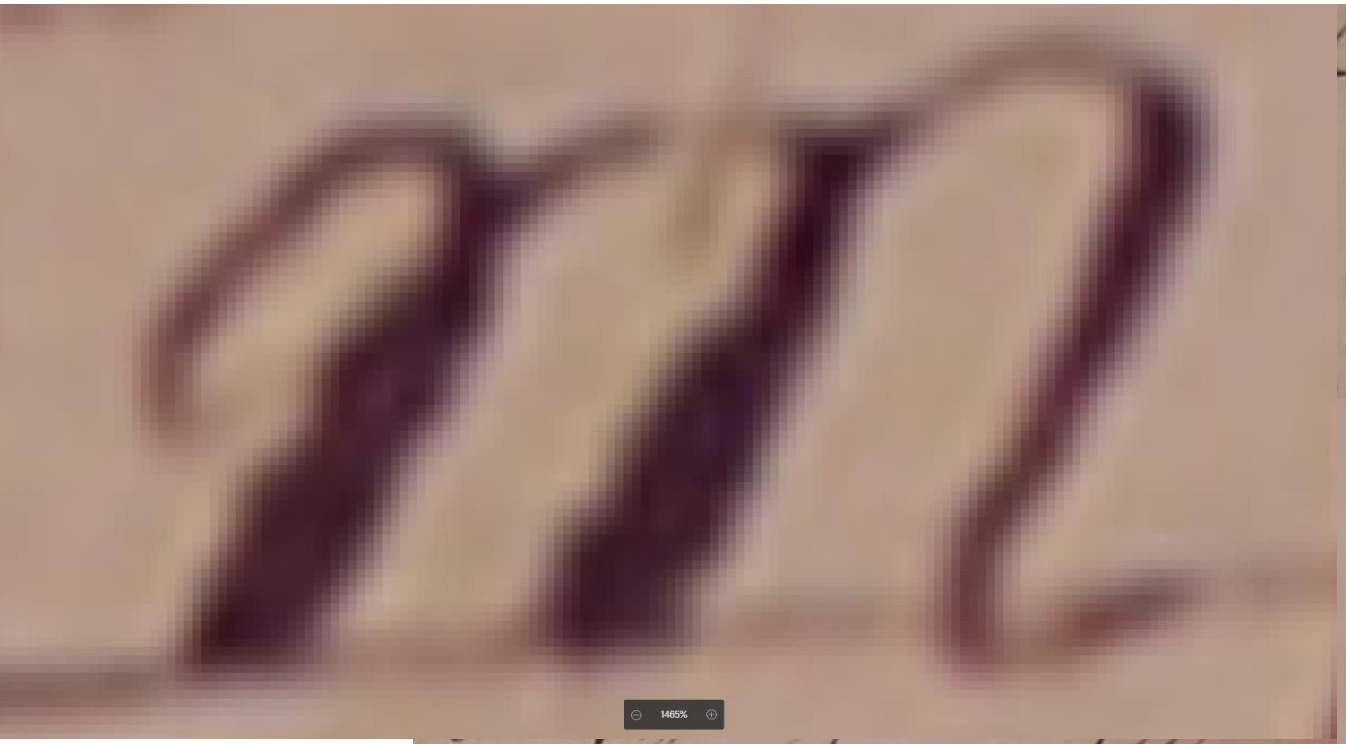
(1) Algorithmic identification of markes, initials and signatures.

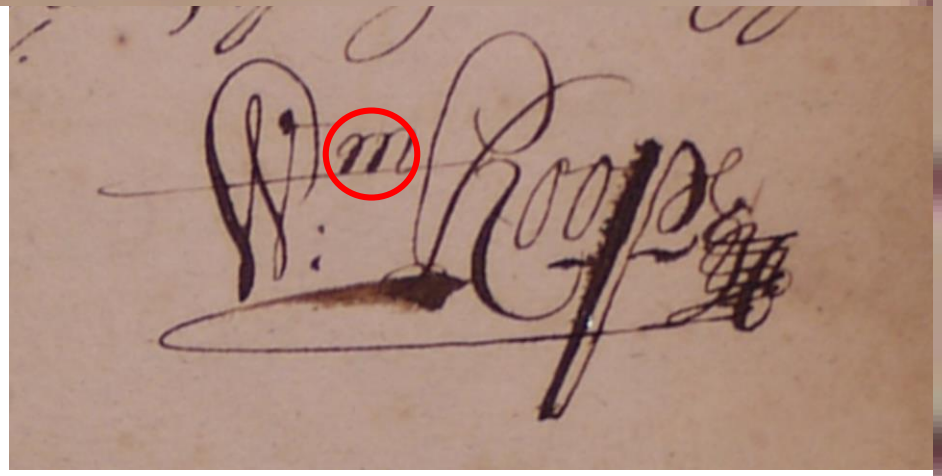
(2) Algorithmic discrimination between degrees of "sophistication" within the three categories of "marke"; "initial(s)", and "signature".

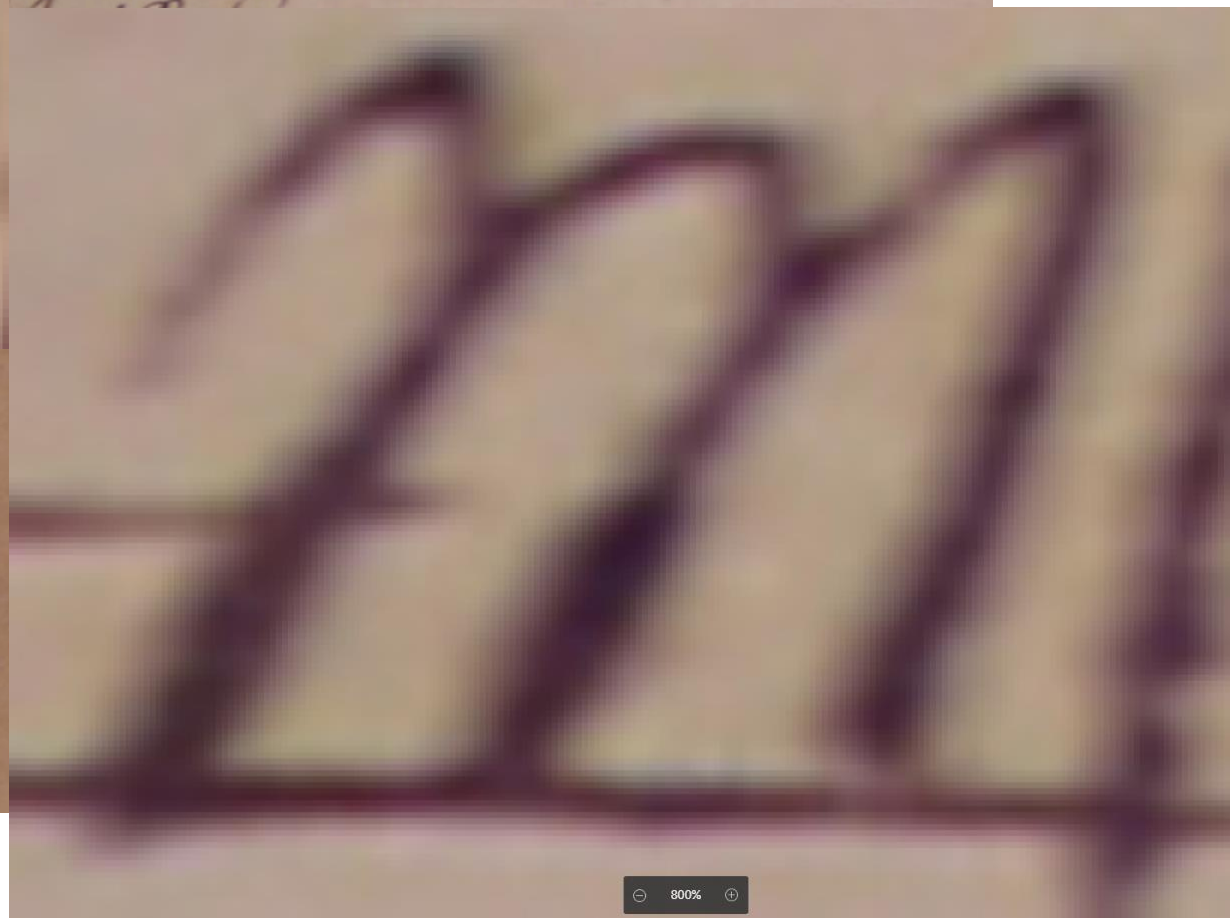
Having proven the concept, we will seek out an image or vision oriented computational laboratory with which to develop a grant funded collaboration to take the work further in 2019 and beyond.

https://www.linkedin.com/pulse/proposed-signs-literacy-kaggle-research-competition-2018-greenstreet/

High pixel definition

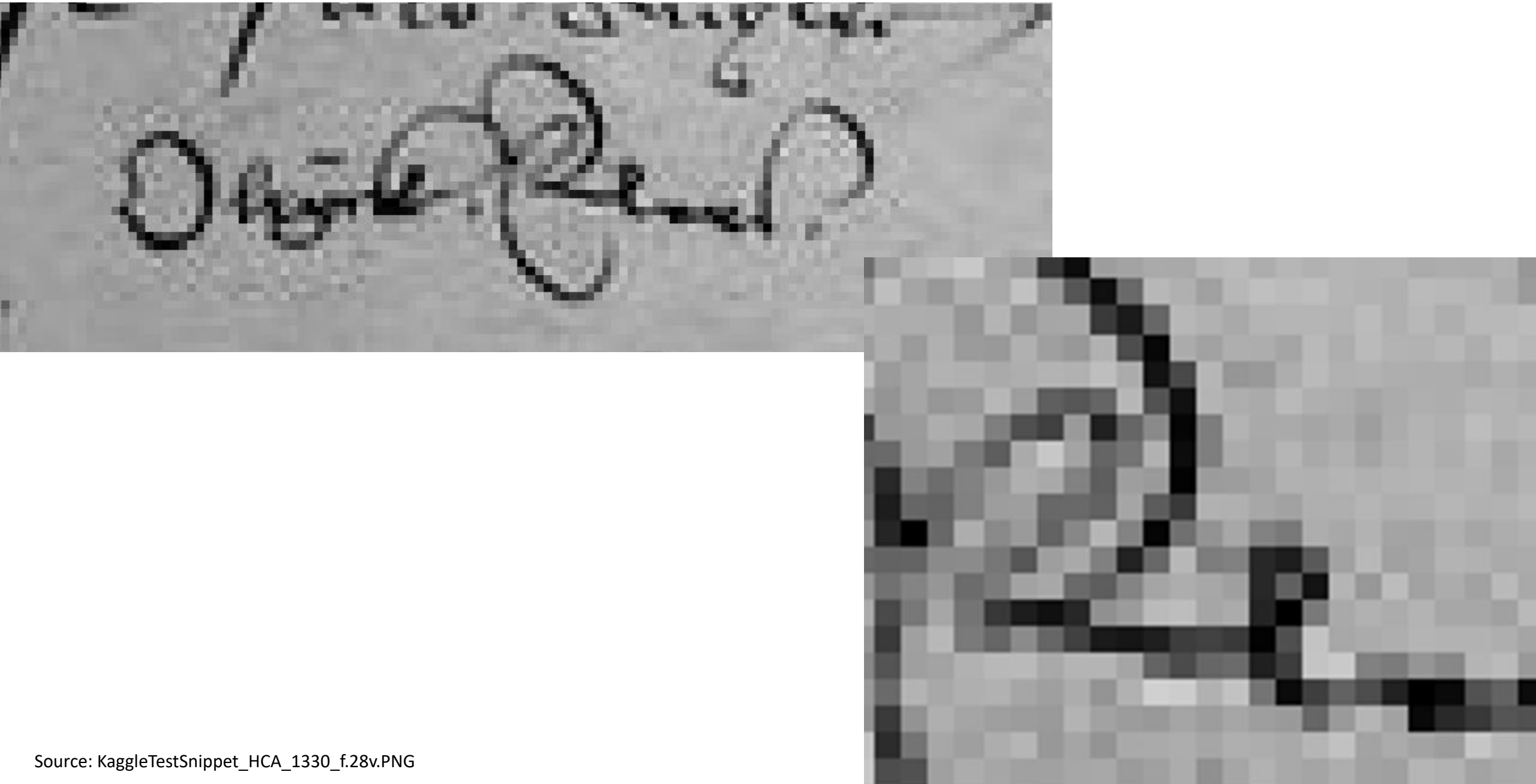






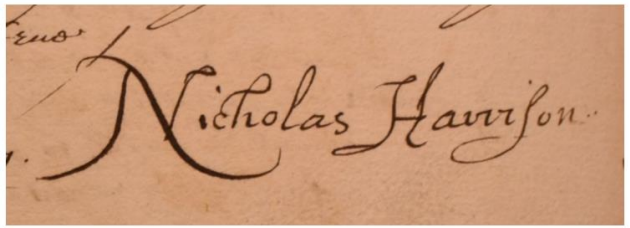
Source: KaggleTestSnippet_HCA_1373_f.199r.PNG,
KaggleTestSnippet_HCA_1373_f.199v_One.PNG

Low pixel definition



Source: KaggleTestSnippet_HCA_1330_f.28v.PNG

Colour analysis – image colour extract PHP, hexadecimal colours

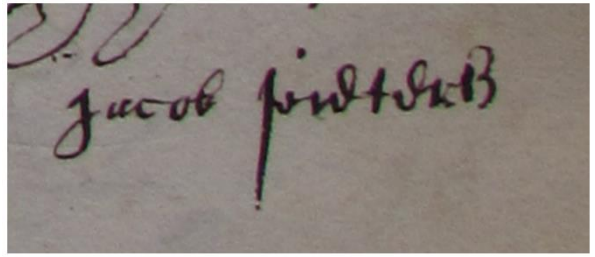


Color	Color Code	Percentage
	#e0a080	0.855975
	#c08060	0.084403
	#806040	0.039371
	#604020	0.013208
	#402020	0.007044

File to upload:
Choose File No file chosen
Number of colors:
10
delta:
32
(1-255)
Reduce Brightness: Yes No
Reduce gradient: Yes No
Run

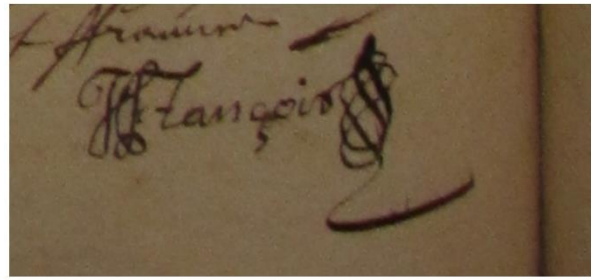
Color	Color Code	Percentage
	#c08060	0.883721
	#806040	0.063798
	#604020	0.048605
	#402000	0.003876

File to upload:
Choose File No file chosen
Number of colors:
10
delta:
32
(1-255)
Reduce Brightness: Yes No
Reduce gradient: Yes No



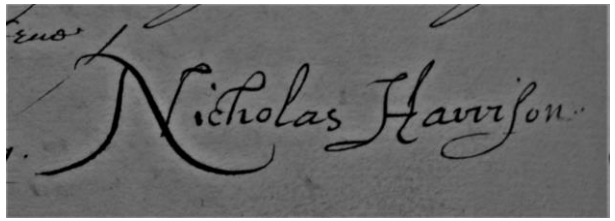
Color	Color Code	Percentage
	#808080	0.969271
	#604040	0.030729

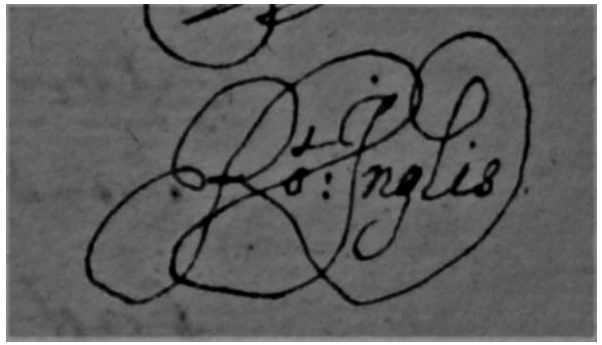
File to upload:
Choose File No file chosen
Number of colors:
10
delta:
32
(1-255)
Reduce Brightness: Yes No
Reduce gradient: Yes No
Run

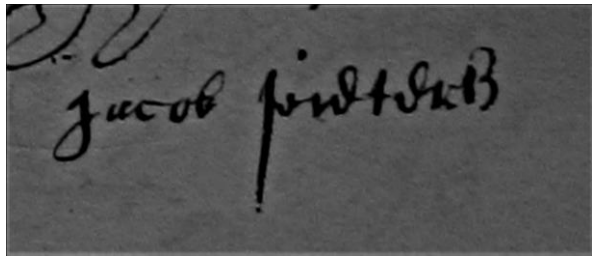


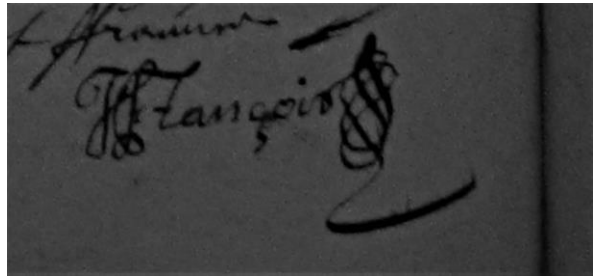
Color	Color Code	Percentage
	#806040	0.806762
	#604020	0.125143
	#402020	0.062000
	#202000	0.006095

File to upload:
Choose File No file chosen
Number of colors:
10
delta:
32
(1-255)
Reduce Brightness: Yes No
Reduce gradient: Yes No
Run



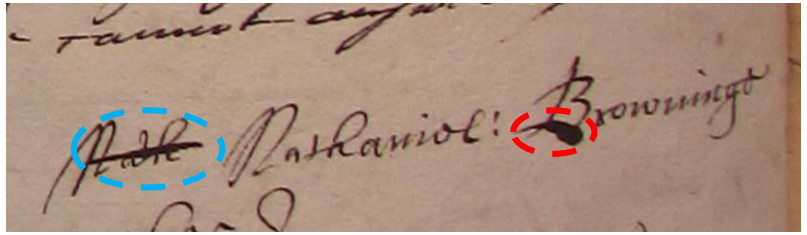


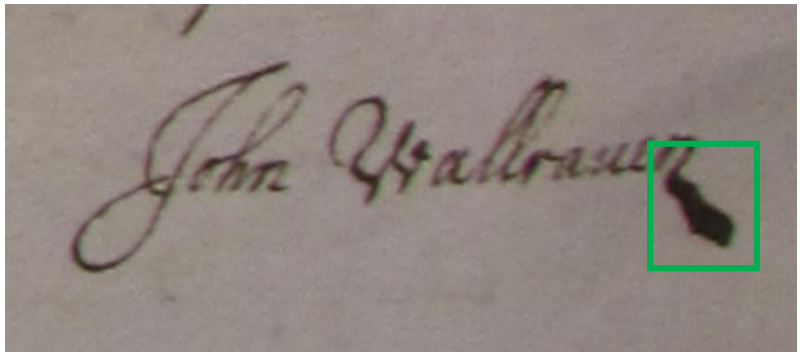


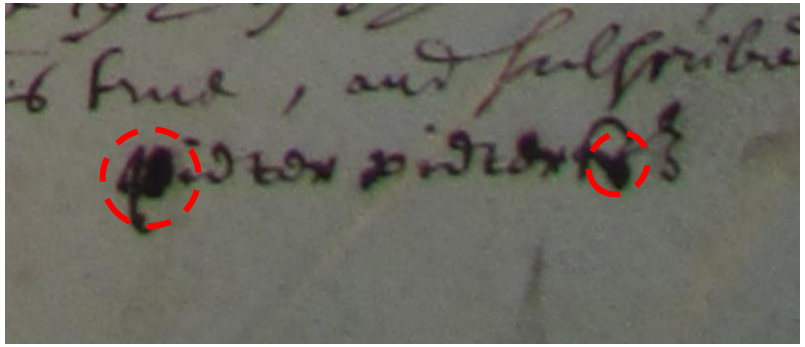


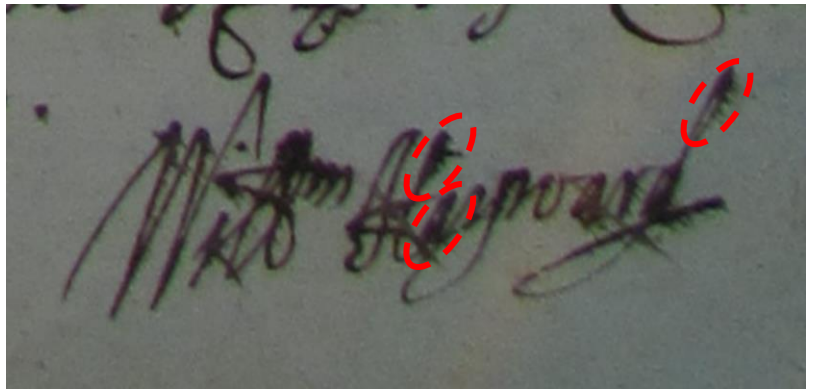
Source: Sample images from SOLM-2018 (KaggleTestSnippet_HCA_1353_f.275v.PNG, KaggleTestSnippet_HCA_1353_f.270v_Two.PNG, KaggleTestSnippet_HCA_1370_f.463r_One.PNG, KaggleTestSnippet_HCA_1368_f.497v.PNG) processed in http://www.coolphptools.com/color_extract#demo; same images reprocessed in Photos SW package, with adjustments set to 0% light, 0% colour, 100% clarity

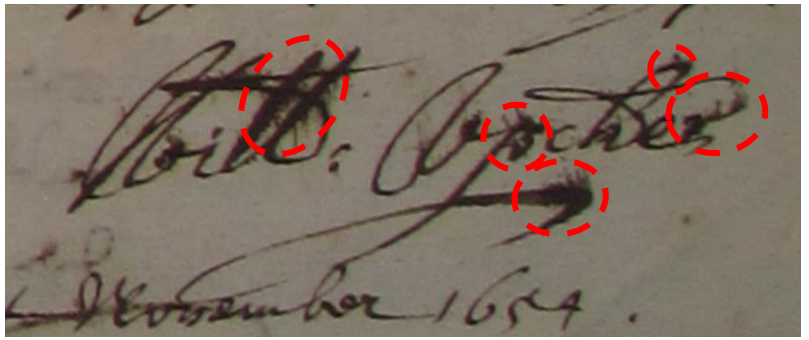
Detection and analysis of blots, smudges, stylistic features, & deletions

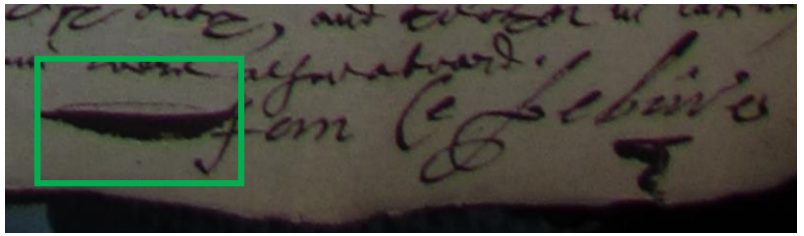




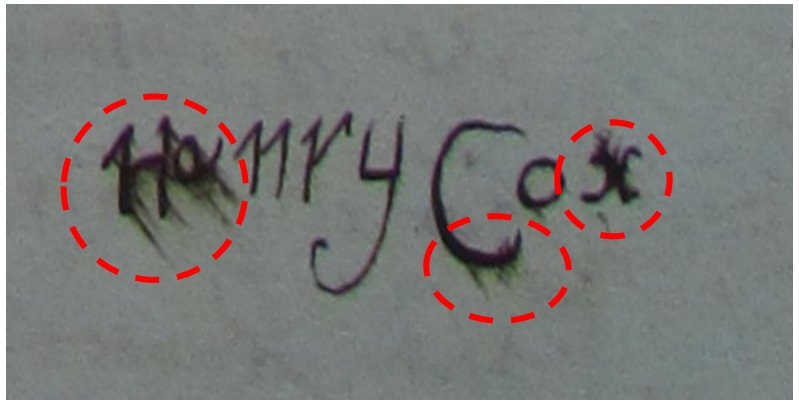


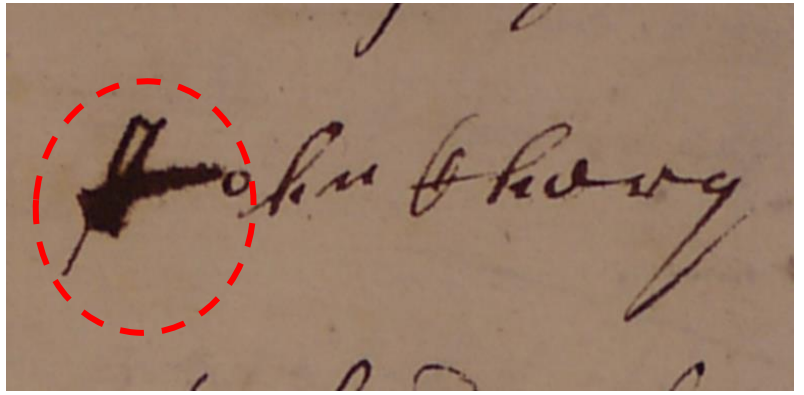






Source: Clockwise from top LH side:
KaggleTestSnippet_HCA_1370_f.387v.PNG,
KaggleTestSnippet_HCA_1370_f.13r.PNG,
KaggleTestSnippet_HCA_1370_f.167r.PNG,
KaggleTestSnippet_HCA_1371_f.456r.PNG,
KaggleTestSnippet_HCA_1370_f.15r.PNG,
KaggleTestSnippet_HCA_1370_f.19r.PNG,
KaggleTestSnippet_HCA_1370_f.41v.PNG,
KaggleTestSnippet_HCA_1370_f.17v.PNG



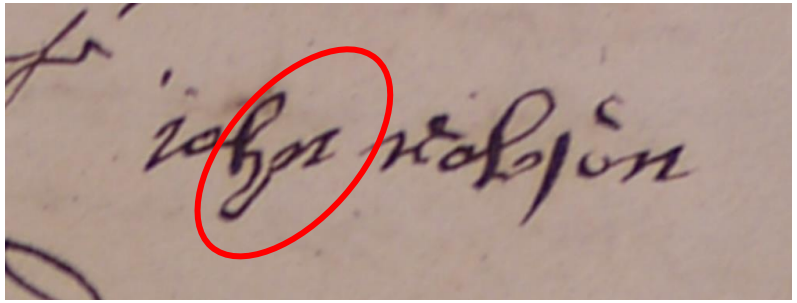


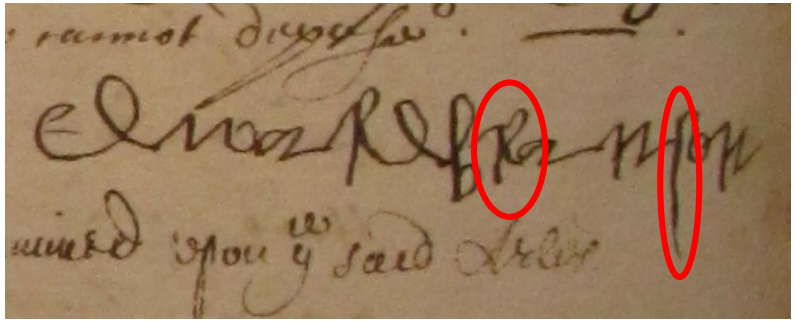
Ink blots or smudges

Stylistic feature or smudge?

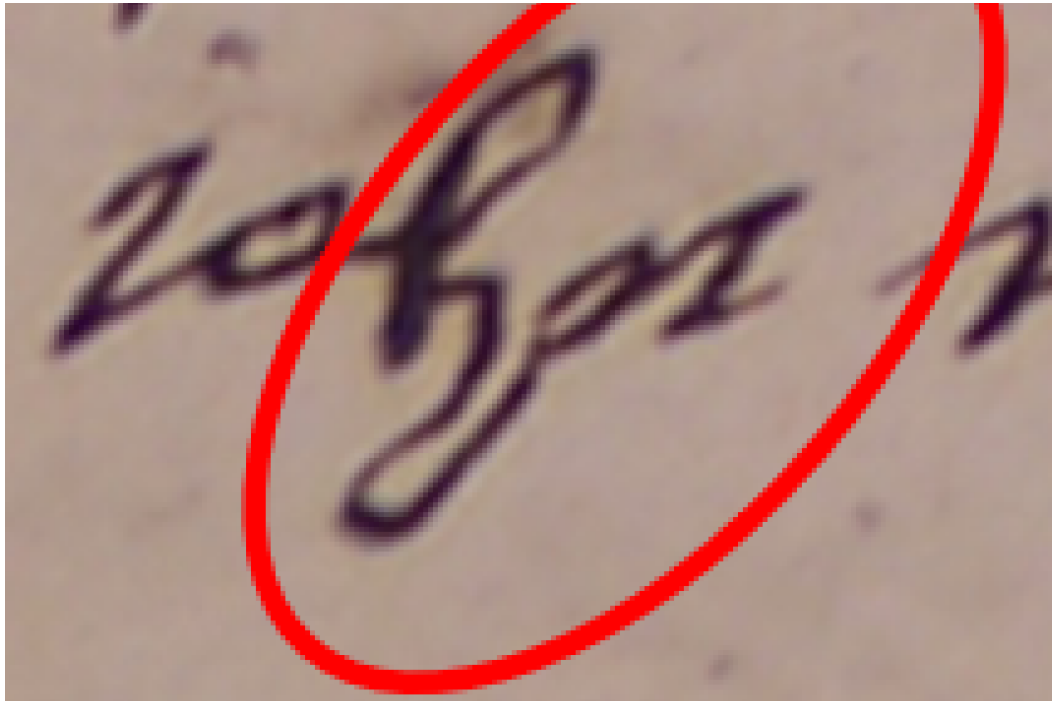
Deletion

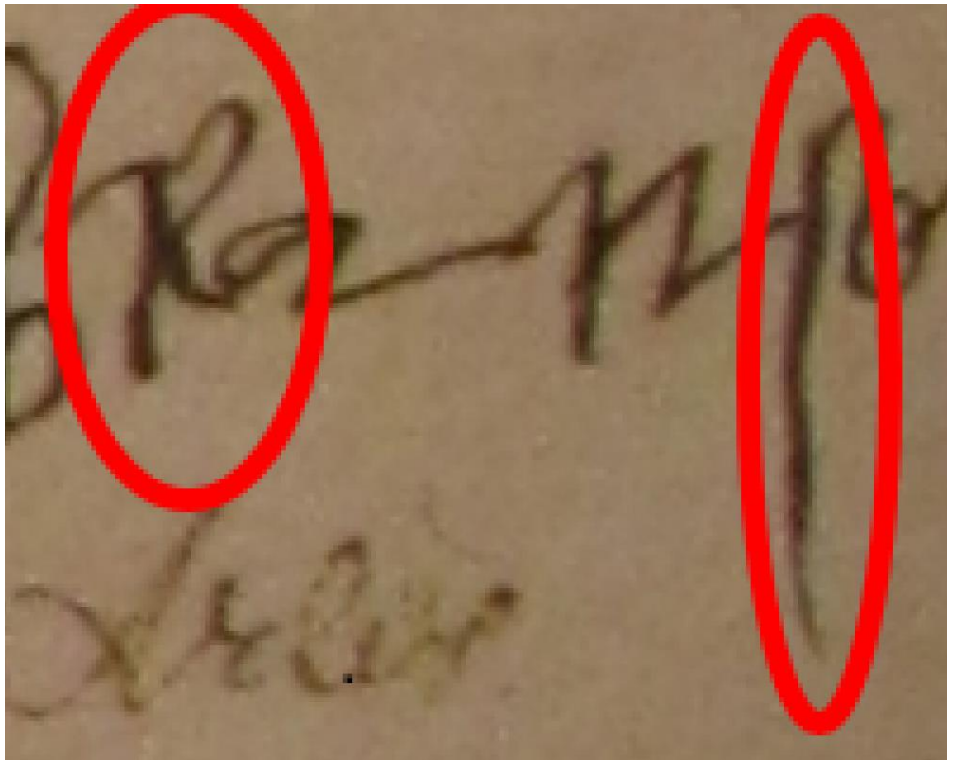
Detection of “shake” in straight and curved lines





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



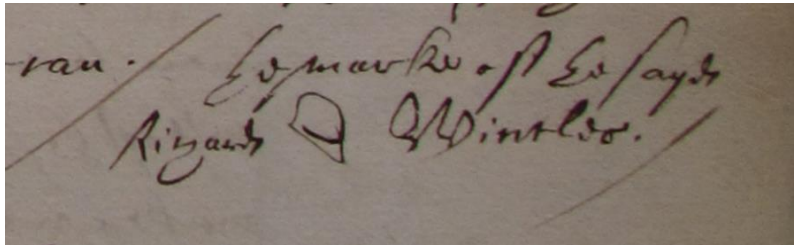


Shaky straight lines and/or loops

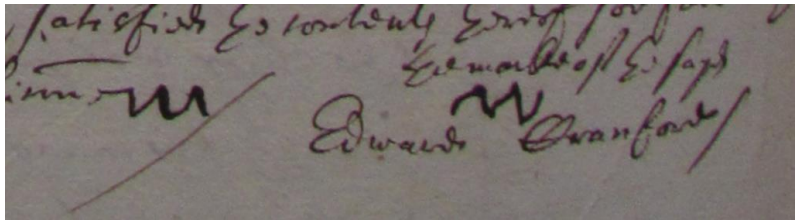
Clockwise, from top LH: KaggleTestSnippet_HCA_1371_f.435v.PNG, KaggleTestSnippet_HCA_1368_f.483v.PNG, KaggleTestSnippet_HCA_1368_f.483v_PIXELS.PNG, KaggleTestSnippet_HCA_1371_f.435v.PNG_PIXELS.PNG

Porters handling coals, whale oil, ginger & corn

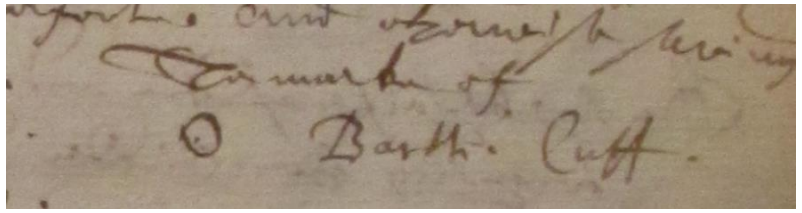
Richard Wincles, thirty-three year old porter, of the parish of Stepney, Middlesex, Dec. 15, 1656; employed as a labourer with fifteen other men to unload coals from the *Imployment* moored near Execution Dock, Wapping, into lighters for fixed rate of 12 *s* per man (HCA 13/70 f.554r)



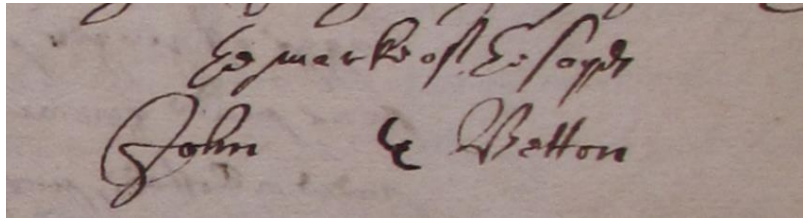
Edward Cranford, forty-four year old coale heaver or porter, of the parish of Stepney, Middlesex, Dec. 15, 1656; employed as a labourer with fifteen other men to unload coals from the *Imployment* moored near Execution Dock, Wapping, into lighters for fixed rate of 12 *s* per man (HCA 13/70 f.555v)



Bartholomew Cuff, sixty year old porter of the Stillyard, of the parish of Allhallowes 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)

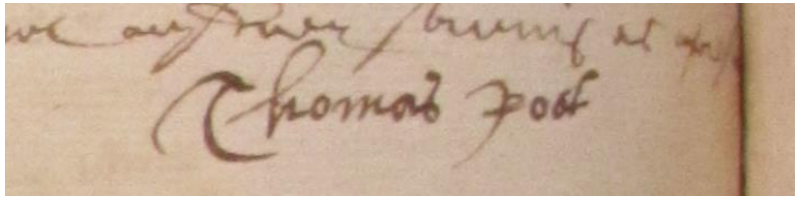


John Betton, fifty-four year old citizen and white baker of London, of the parish of Saint Buttolph Algate, London, Jul. 31, 1655; self-described as a porter employed by the Commissioners for Prize Goods to deliver ginger from a warehouse at Ralphes Key (HCA 13/70 f.449r)

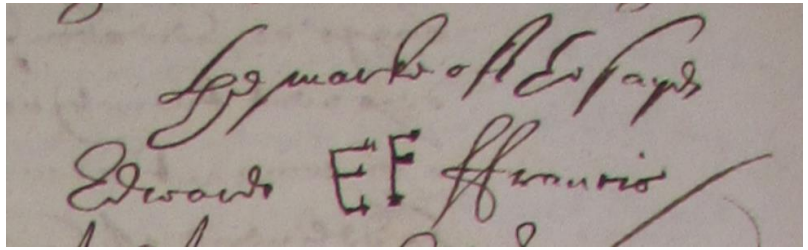


Edward Sherwin, fifty-six year old cittizen and leatherseller, of the parish of Little Allhallowes, London, Jul. 31, 1655; self-described as a porter employed by the Commissioners for Prize Goods to deliver ginger from a warehouse at Ralphes Key (HCA 13/70 f.449v)

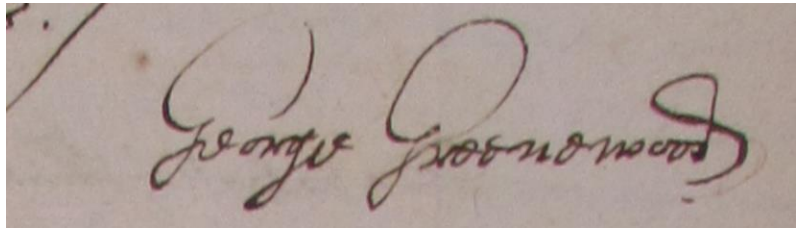
Thomas Roots, twenty-nine year old porter, of the parish of Greate Allhallowes, London, May 15, 1658; assisted in the landing of whale oil from lighters at the Stillyard Key, as one of the Stillyard porters, and loading them away into a warehouse (HCA 13/72 f.330v)



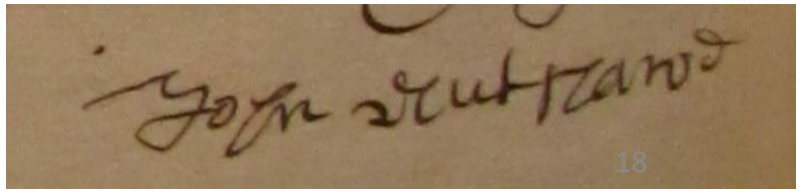
Edward ffrancis, citizen and merchant taylor of London, of the parish of Saint Olave in Southwarke, Jul. 31, 1655; self-described as a porter employed by the Commissioners for Prize Goods to deliver ginger from a warehouse at Ralphes Key (HCA 13/70 f.450v)



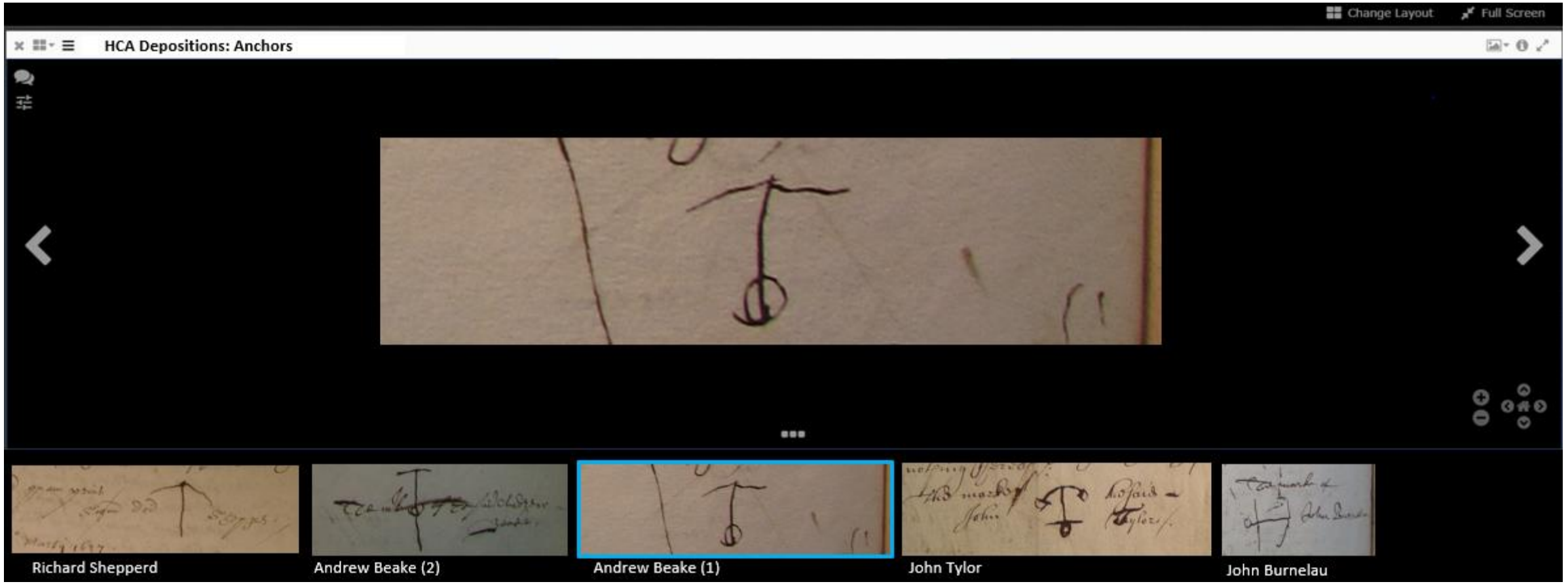
George Greenwood, thirty year old citizen and vintner of London, of the parish of Saint Buttolph Bishopsgate, London, Jul. 31, 1655; self-described as a porter employed by the Commissioners for Prize Goods to deliver ginger from a warehouse at Ralphes Key (HCA 13/70 f.454r)



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

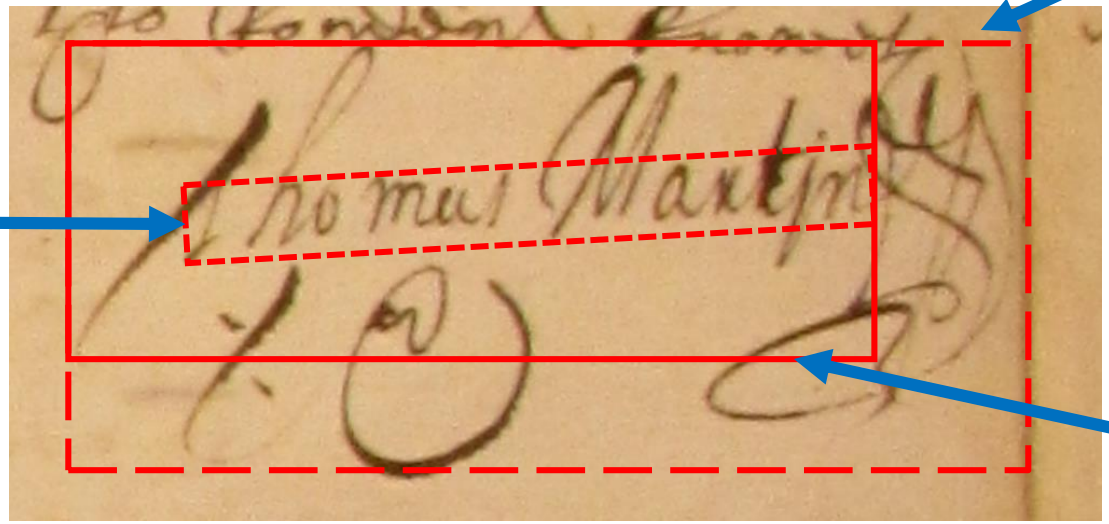


Mockup of a IIIF manifest in Mirador viewer, using http://projectmirador.org/demo/ ;
http://www.marinelives.org/wiki/HCA_13/70_f.252v_Annotate

Boundary boxes marking the visual geometry of a signature

Outside boundary box, including flourish

Inside boundary box, excluding uppers and downers

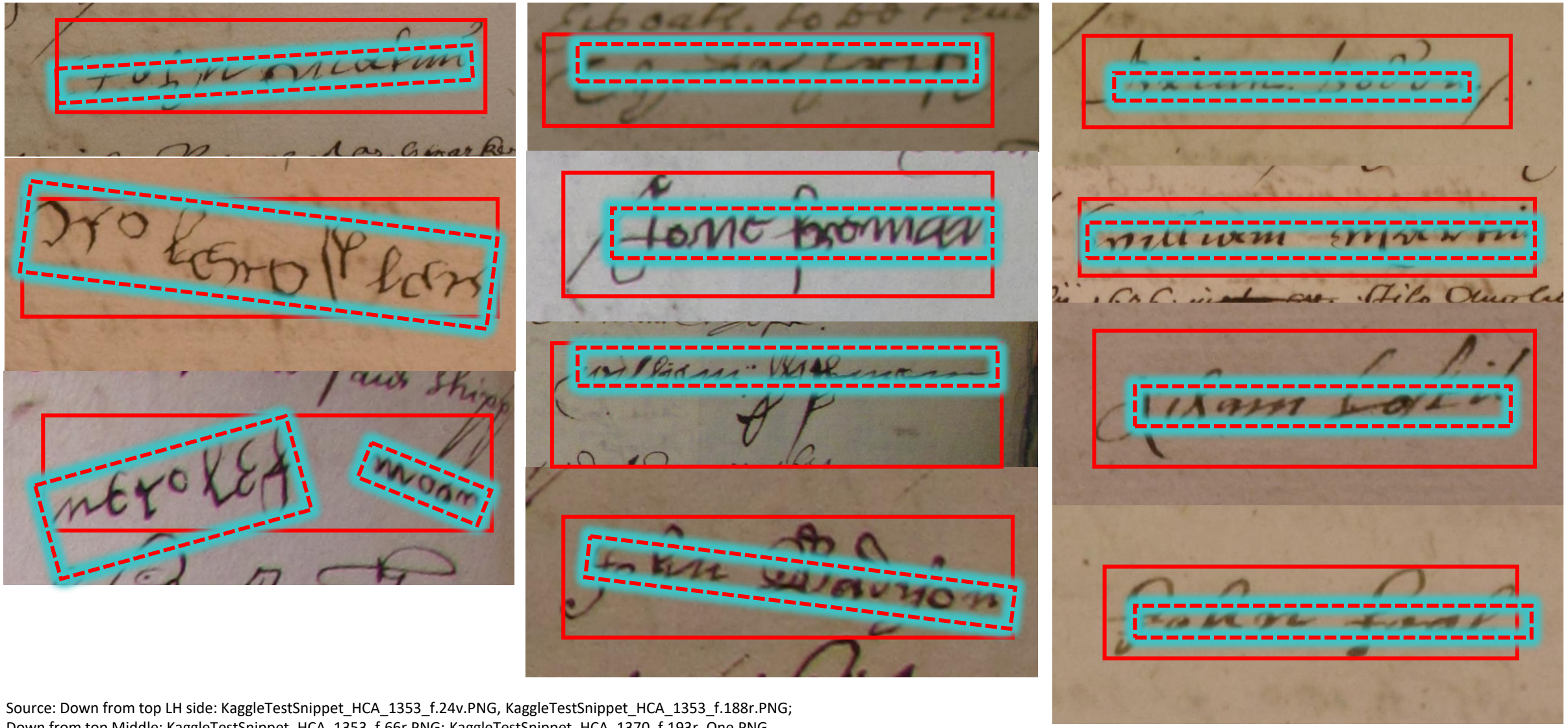


Middle boundary box, including all letters, but excluding flourish

Statistics

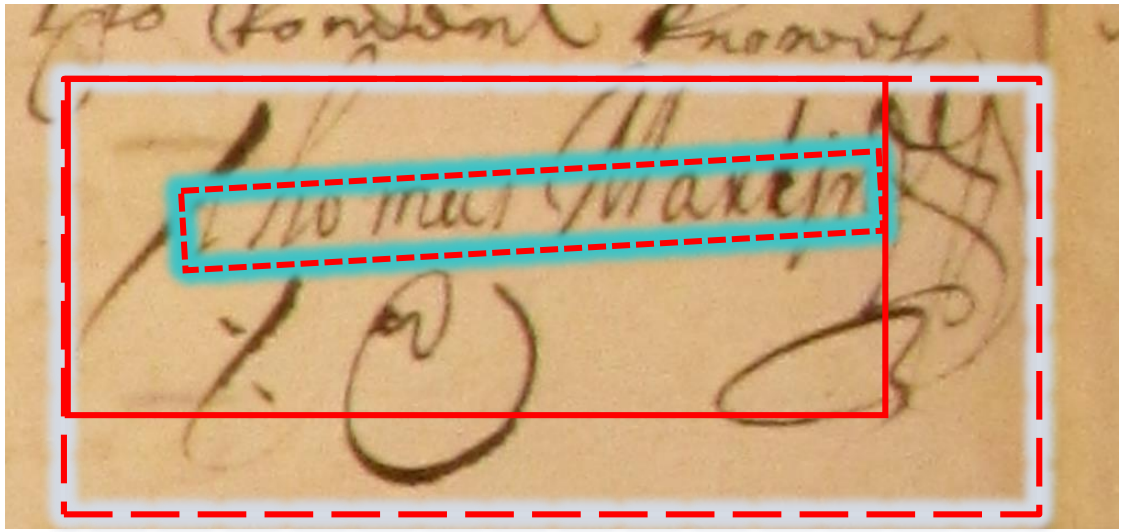
Inside boundary box: 9.0 x 1.1
Middle boundary box: 9.75 x 4.25
Outside boundary box: 12.75 x 5.75
Rotation from horizontal: ca. 340 degrees

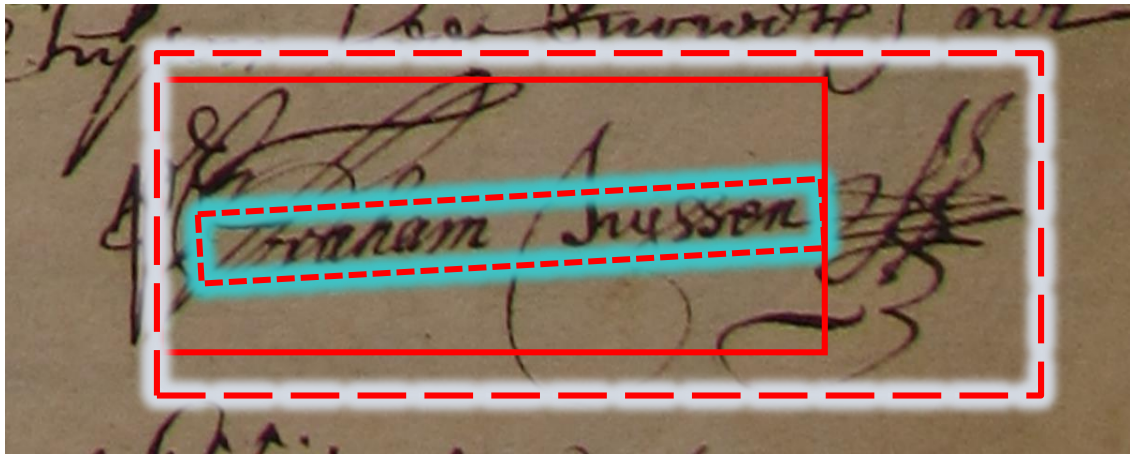
Simple signatures, no flourishes

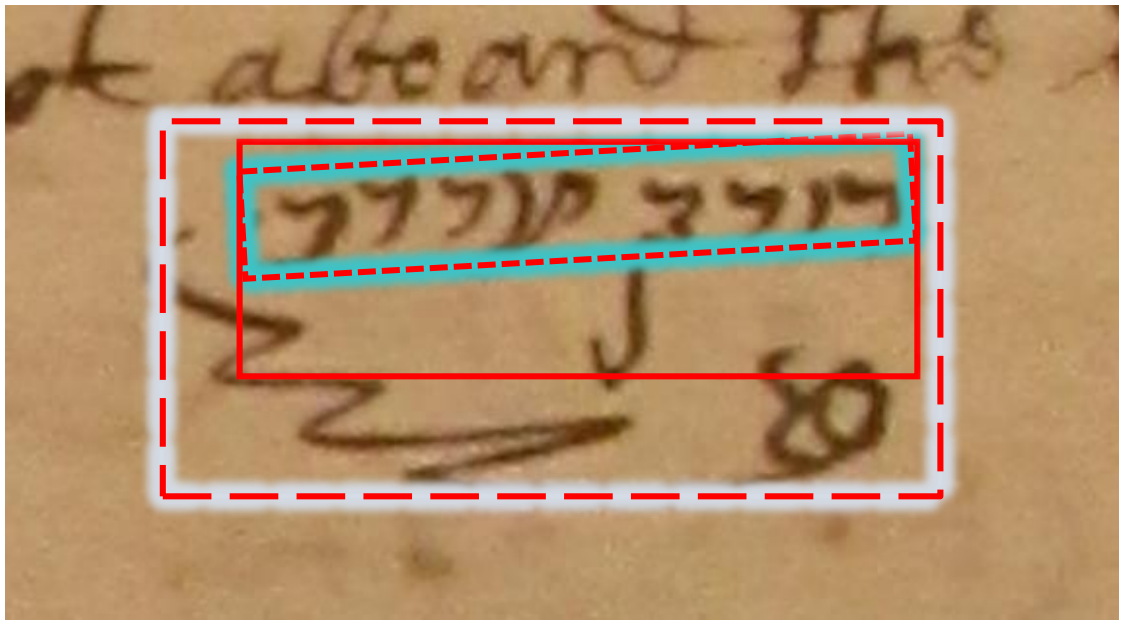


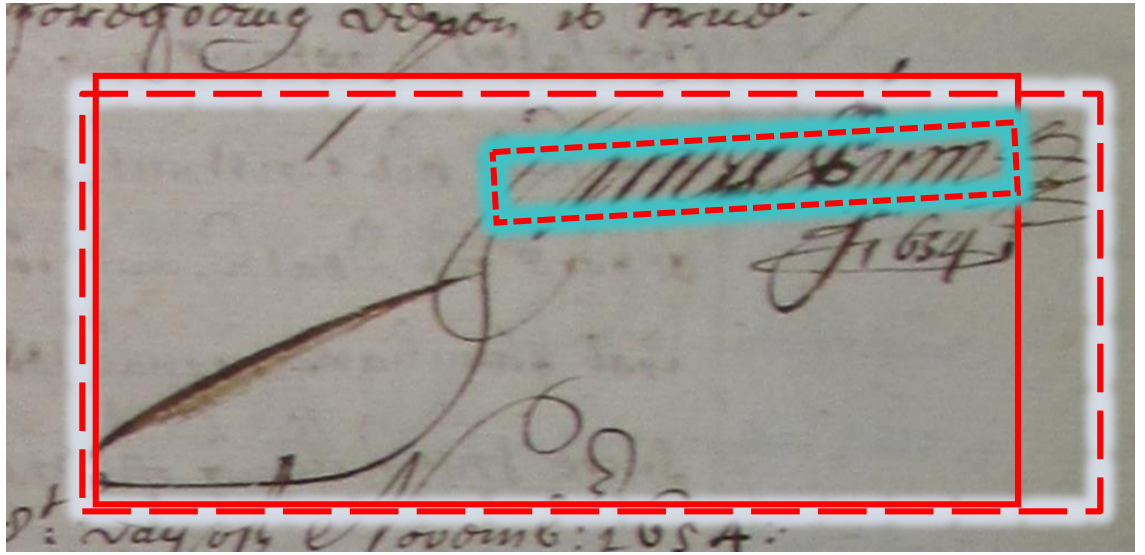
Source: Down from top LH side: KaggleTestSnippet_HCA_1353_f.24v.PNG, KaggleTestSnippet_HCA_1353_f.188r.PNG;
Down from top Middle: KaggleTestSnippet_HCA_1353_f.66r.PNG; KaggleTestSnippet_HCA_1370_f.193r_One.PNG,
KaggleTestSnippet_HCA_1370_f.203r.PNG, KaggleTestSnippet_HCA_1370_f.218r.PNG
Down from top 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





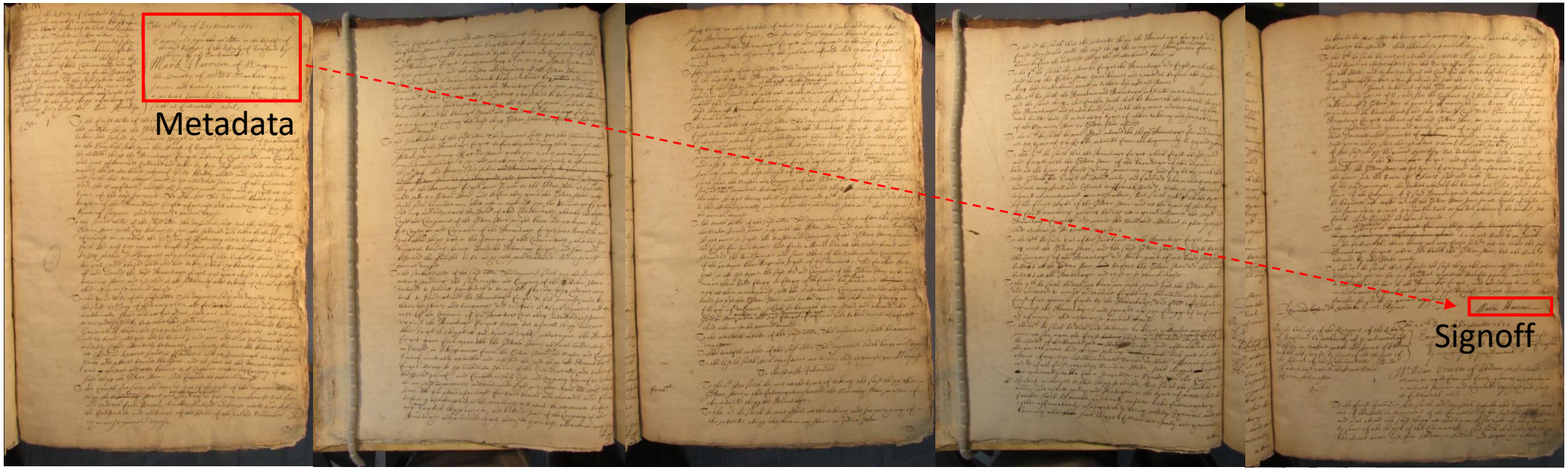




Source: Clockwise from top LH side: KaggleTestSnippet_HCA_1368_f.34v.PNG, KaggleTestSnippet_HCA_1370_f.366r.PNG, KaggleTestSnippet_HCA_1370_f.134r.PNG, KaggleTestSnippet_HCA_1368_f.58r.PNG

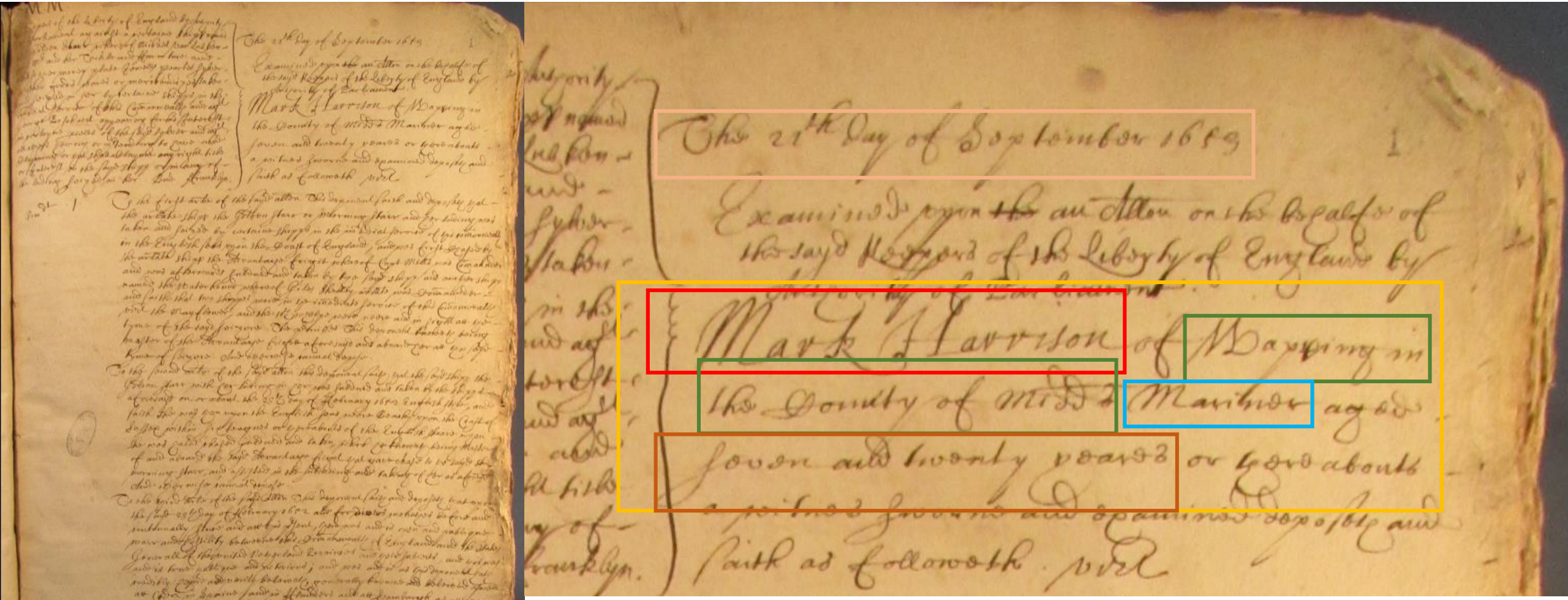
Legal deposition

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



http://www.marinelives.org/wiki/HCA_13/68_f.1r_Annotate; http://www.marinelives.org/wiki/HCA_13/68_f.1v_Annotate; http://www.marinelives.org/wiki/HCA_13/68_f.2r_Annotate ; http://www.marinelives.org/wiki/HCA_13/68_f.2v_Annotate; http://www.marinelives.org/wiki/HCA_13/68_f.3r_Annotate

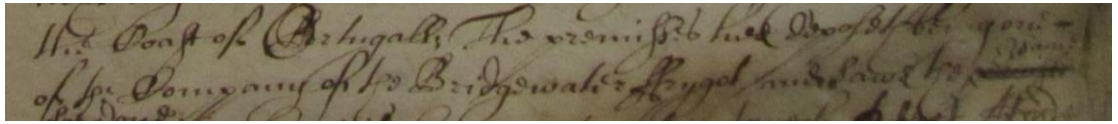
Machine based recognition of metadata



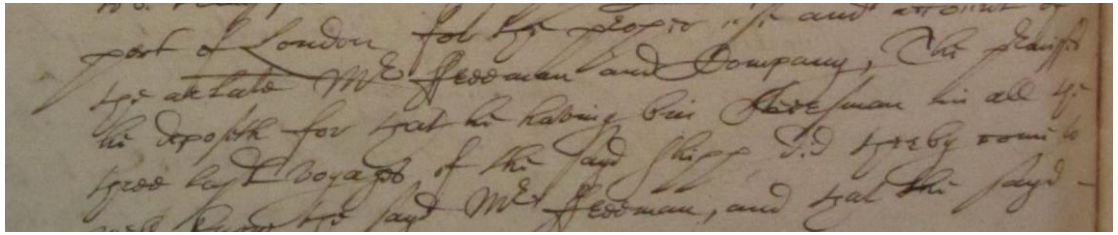
http://www.marinelives.org/wiki/HCA_13/68_f.1r_Annotate;
http://www.marinelives.org/wiki/Mark_Harrison

Can we use key word spotting to excavate raw metadata?

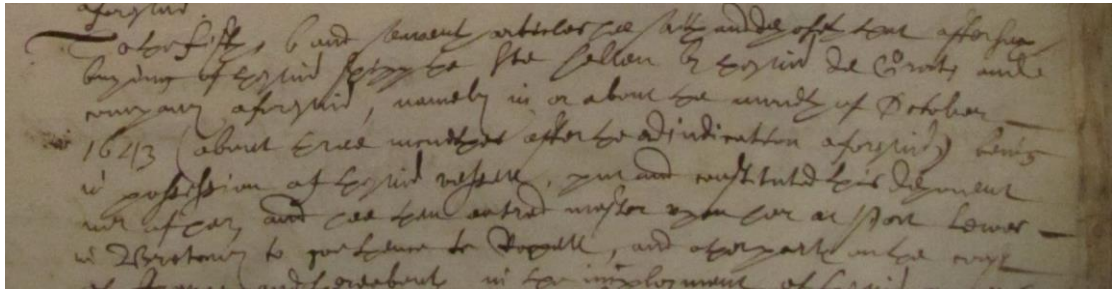
LANGUAGE DENOTING OCCUPATION



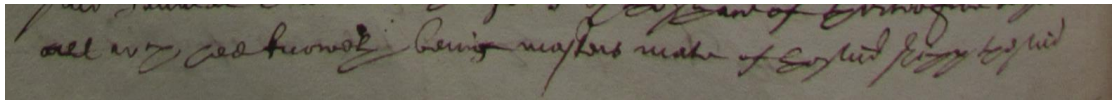
"The premisses hee deposeth being one of the company of the *Bridgewater ffrygott*, and sawe the same soe done" [HCA 13/72 f.90r] [CONCLUSION: One of the company]



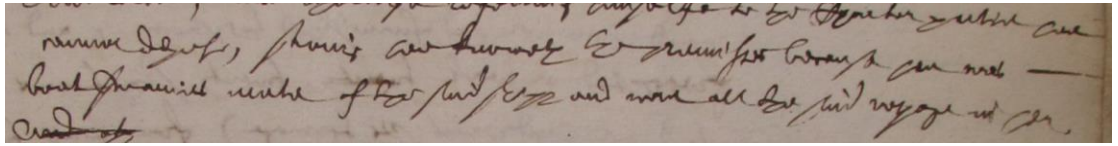
"The premisses he deposeth for that he the deponent was not onely for the voyage arlate wherein she was stranded, but in two former voyages stiersman of the sayd ship" [HCA 13/72 f.90v] [CONCLUSION: Steersman]



"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 adiudication aforesaid) being in possession of the said vessell, 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]



"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]

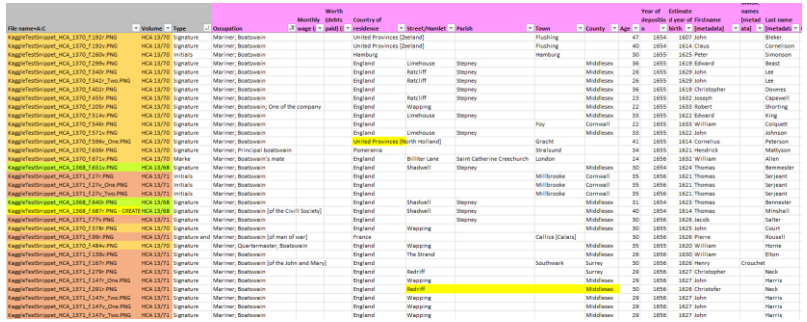
"The premisses he deposeth for that he the deponent was not onely for the voyage arlate wherein she was stranded, but in two former voyages stiersman of the sayd ship" [HCA 13/72 f.90v] [CONCLUSION: Steersman]

"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 adiudication aforesaid) being in possession of the said vessell, 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]

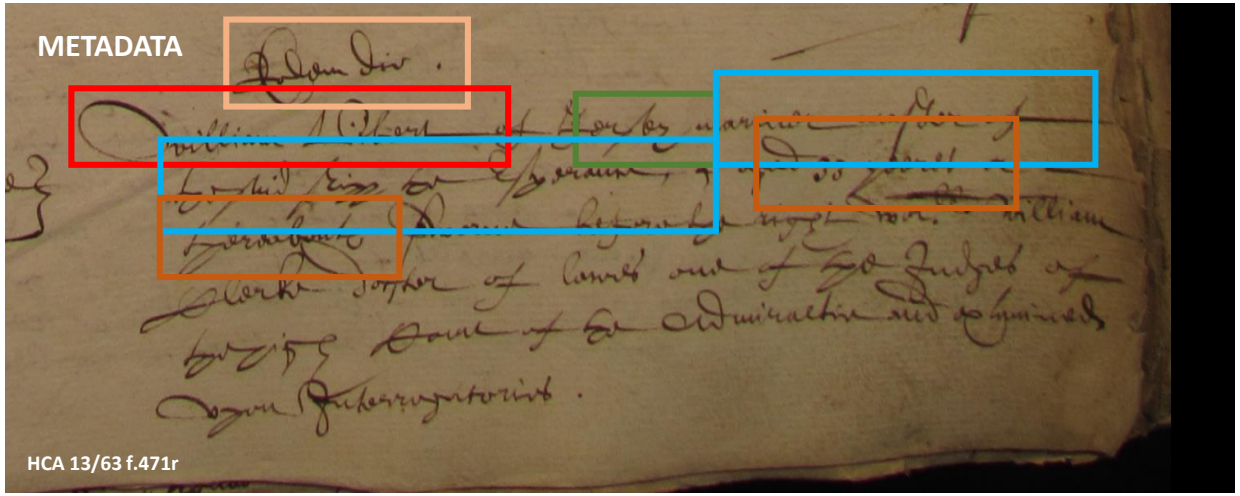
"all which hee knoweth being masters mate of the said shipp the said voyage" [HCA 13/70 f.669v] [CONCLUSION: Master's mate]

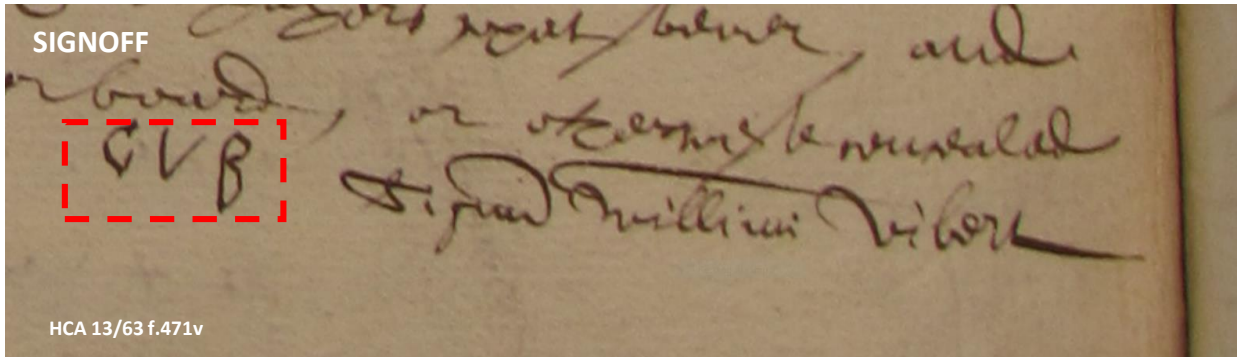
"the premisses because hee was boatswaines mate of the said shipp and went all the said voyage in her" [HCA 13/70 f.671r] [CONCLUSION: Boatswain's mate]

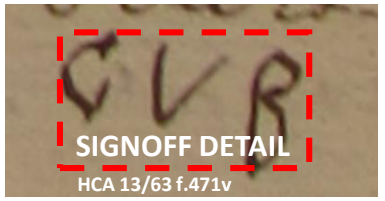
KaggleTestSnippet_HCA_1370_f.546r.PNG	HCA 13/70	Signature	Mariner; Boatswain	
KaggleTestSnippet_HCA_1370_f.571v.PNG	HCA 13/70	Signature	Mariner; Boatswain	
KaggleTestSnippet_HCA_1370_f.596v_One.PNG	HCA 13/70	Signature	Mariner; Boatswain	
KaggleTestSnippet_HCA_1370_f.636r.PNG	HCA 13/70	Signature	Mariner; Principal boatswain	
KaggleTestSnippet_HCA_1370_f.671v.PNG	HCA 13/70	Marke	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.687r.PNG - CREATE	HCA 13/68	Signature	Mariner; Boatswain [of the Civill 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





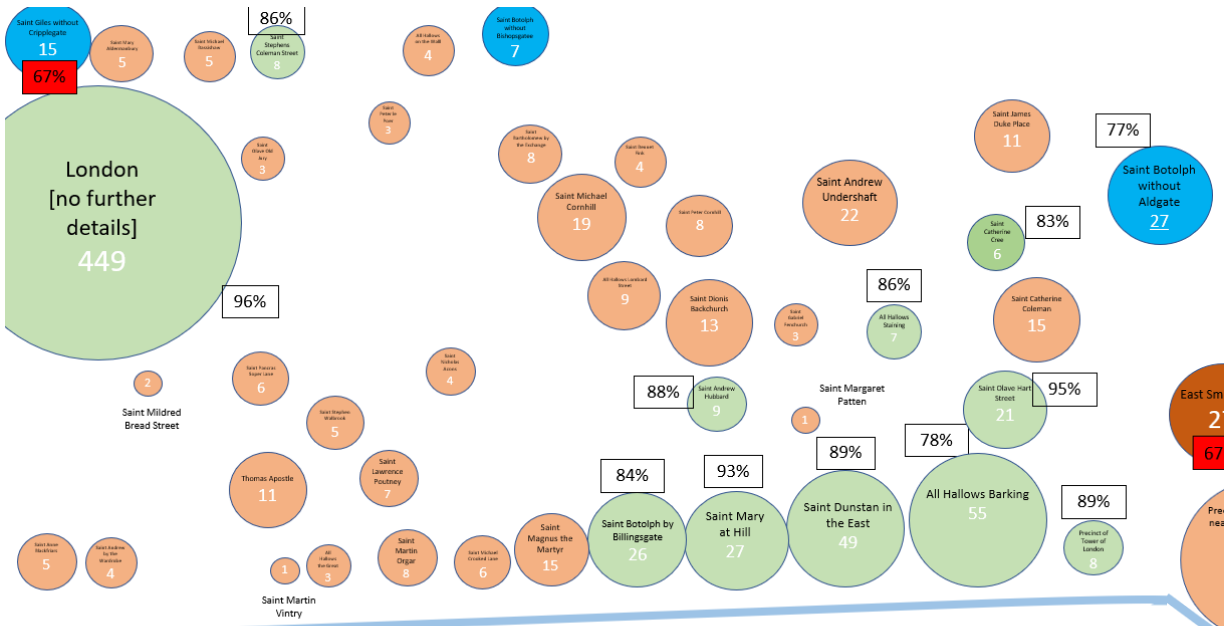


Date
Occupation
Name
Age
Residence
Signoff

Source: http://www.marinelives.org/wiki/HCA_13/63_f.471r_Annotate; http://www.marinelives.org/wiki/HCA_13/63_f.471v_Annotate

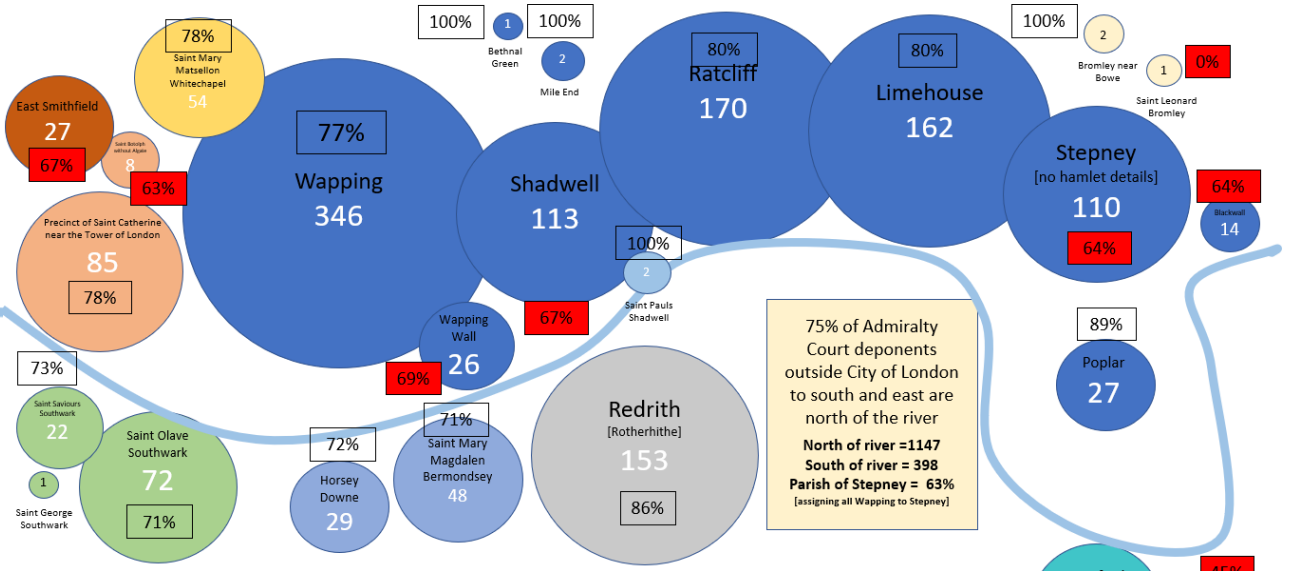
Some early results

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



n = 932 [Partial parish based data set, including all signoffs labelled only "London", without parish data]. Data set as of 03/09/2018

SOLM-2018 data set: Number of signoffs by parish in city of London [partial]



SOLM-2018 data set: Number of signoffs by hamlet & parish to east & south east of city of London (1592-1683)

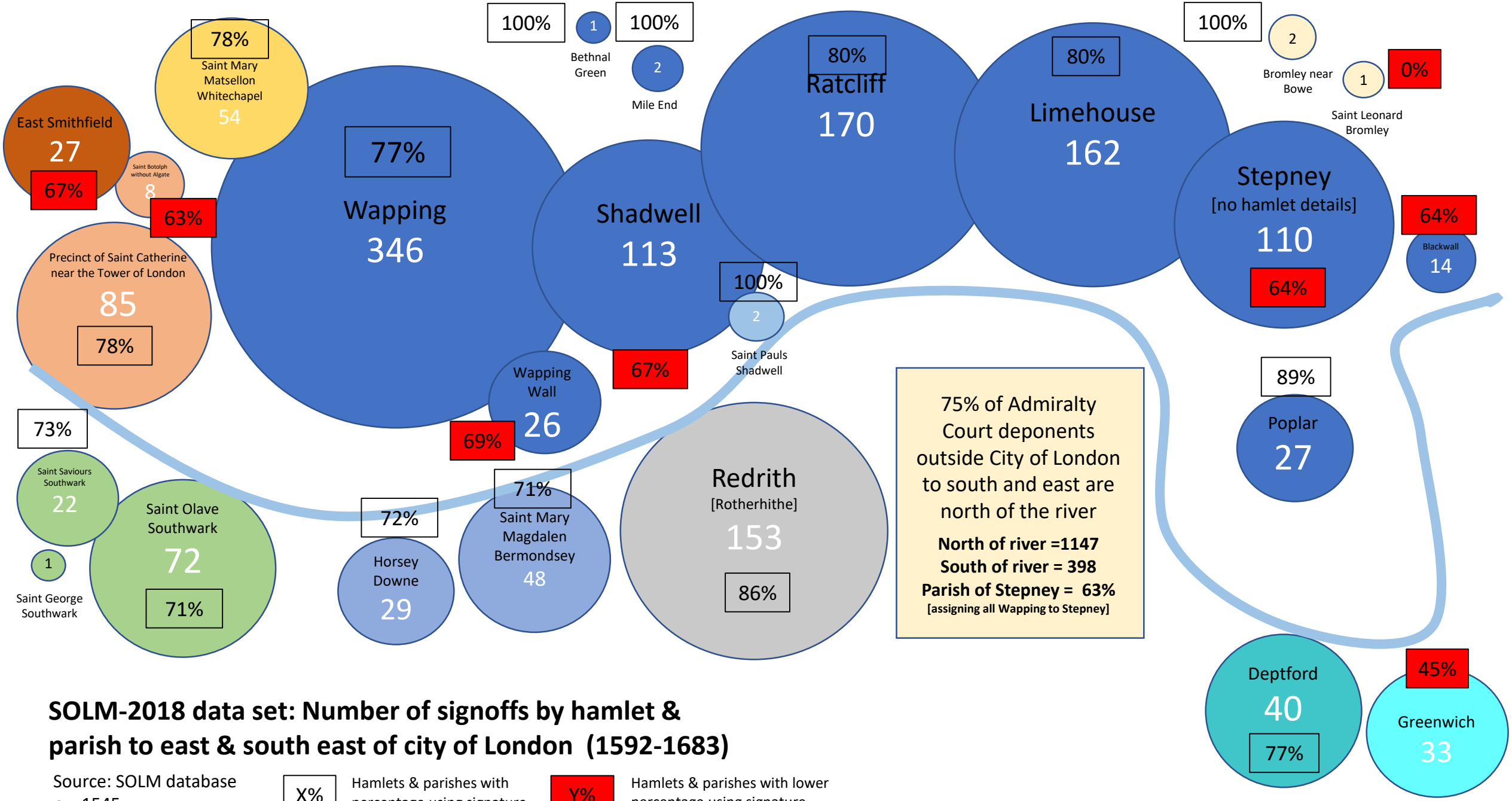
Source: SOLM database
n = 1545
Data set as of 03/09/2018

X% Hamlets & parishes with percentage using signature

Y% Hamlets & parishes with lower percentage using signature

Some early results 31

Source: SOLM-2018 database



SOLM-2018 data set: Number of signoffs by hamlet & parish to east & south east of city of London (1592-1683)

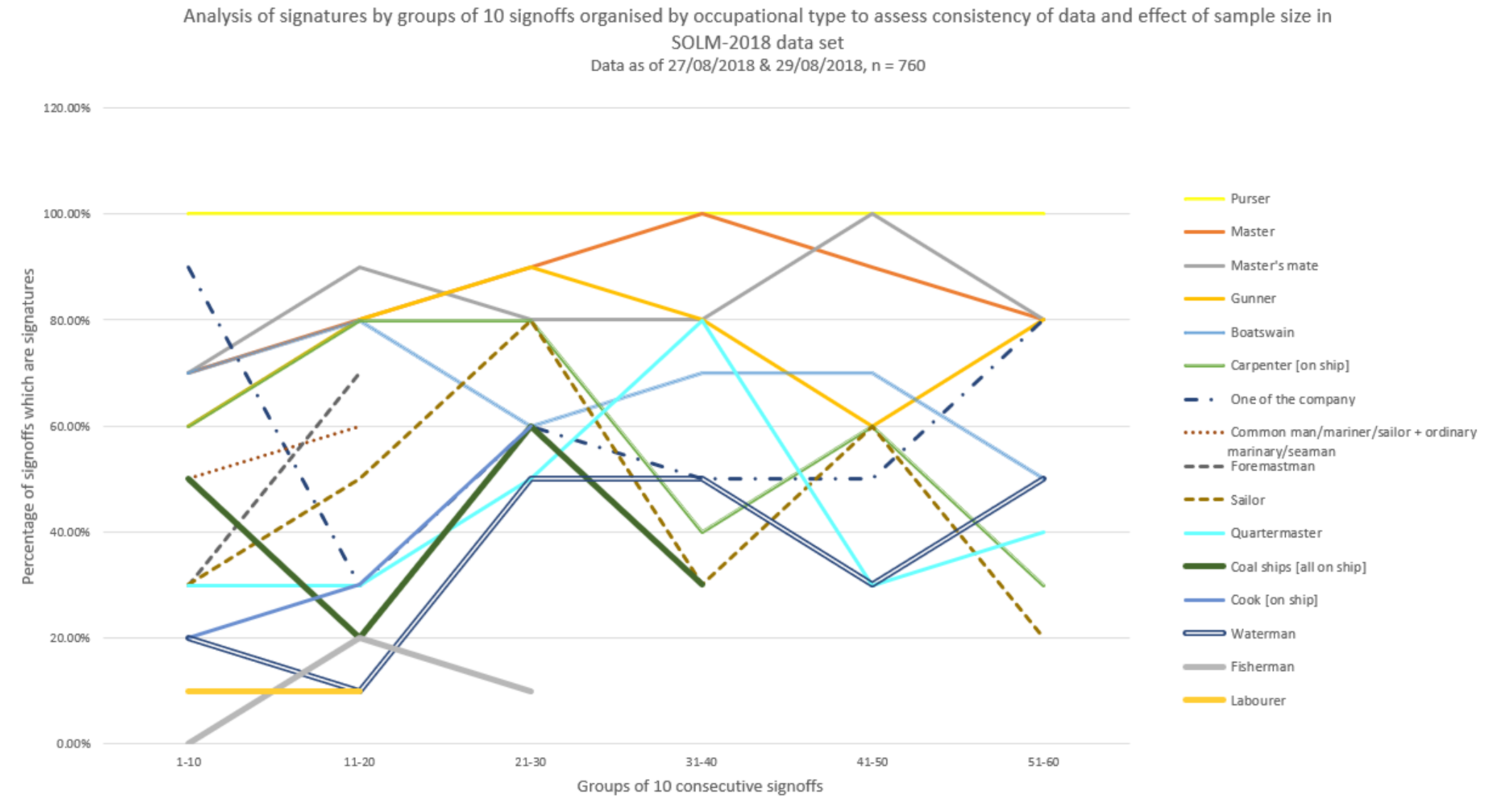
Source: SOLM database
n = 1545
Data set as of 03/09/2018

X% Hamlets & parishes with percentage using signature

Y% Hamlets & parishes with lower percentage using signature

Some early results

Some early results

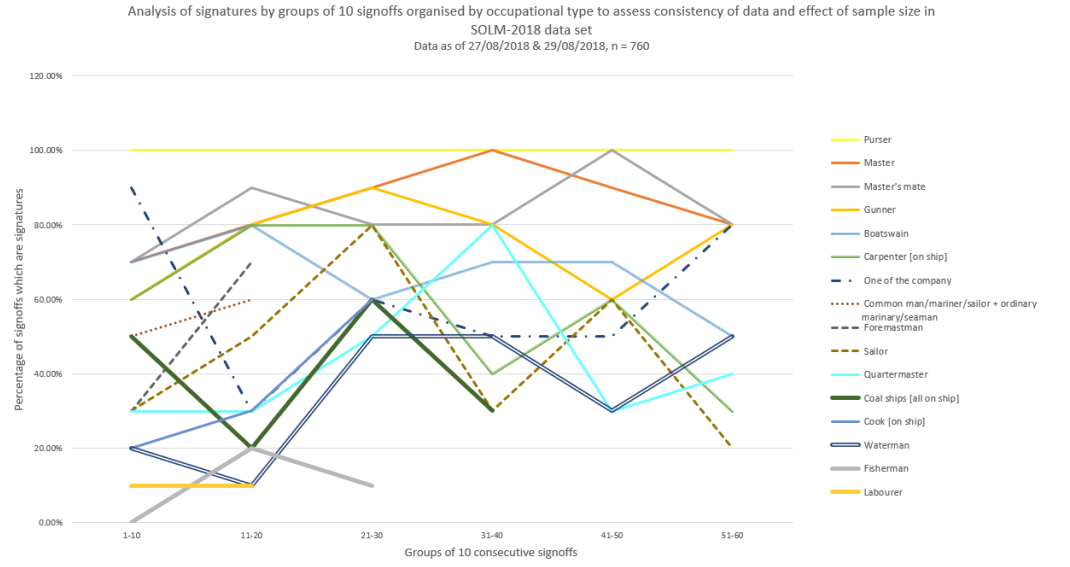


Source: SOLM-2018 database

Some early results

Table 20.0: Signoff composition by Occupational group
PURSERS, MASTERS, MASTER'S MATES, GUNNERS/GUNNERS' MATES, BOATSWAINS, , CARPENTERS (ON SHIPS), ONE OF THE COMPANY, FOREMASTMEN, SAILORS, COOKS [ON SHIPS], WATERMEN, FISHERMEN AND LABOURERS
Data as of 27/08/2018 & 29/08/2018

	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count
Group	Purser	Master	Master's mate	Gunner	Boatswain	Carpenter [on ship]	One of the company	Common man/mariner /sailor + ordinary marinary/sea man	Foremastman	Sailor	Quartermaster	Coal ships [all on ship]	Cook [on ship]	Waterman	Fisherman	Labourer
1-10	100.00%	70.00%	70.00%	60.00%	70.00%	60.00%	90.00%	50.00%	30.00%	30.00%	30.00%	50.00%	20.00%	20.00%	0.00%	10.00%
11-20	100.00%	80.00%	90.00%	80.00%	80.00%	80.00%	30.00%	60.00%	70.00%	50.00%	30.00%	20.00%	30.00%	10.00%	20.00%	10.00%
21-30	100.00%	90.00%	80.00%	90.00%	60.00%	80.00%	60.00%			80.00%	50.00%	60.00%	60.00%	50.00%	10.00%	
31-40	100.00%	100.00%	80.00%	80.00%	70.00%	40.00%	50.00%			30.00%	80.00%	30.00%		50.00%		
41-50	100.00%	90.00%	100.00%	60.00%	70.00%	60.00%	50.00%			60.00%	30.00%			30.00%		
51-60	100.00%	80.00%	80.00%	80.00%	50.00%	30.00%	80.00%			20.00%	40.00%			50.00%		
STATISTICAL MEASURES																
Standard deviation of population	0.00%	9.57%	9.43%	11.18%	9.43%	18.63%	20.00%	5.00%	20.00%	20.62%	17.95%	15.81%	17.00%	16.07%	8.16%	0.00%
Median	100.00%	85.00%	80.00%	80.00%	70.00%	60.00%	55.00%	55.00%	50.00%	40.00%	35.00%	40.00%	30.00%	40.00%	10.00%	10.00%
Average	100.00%	85.00%	83.33%	75.00%	66.67%	58.33%	60.00%	55.00%	50.00%	45.00%	43.33%	40.00%	36.67%	35.00%	10.00%	10.00%



Source: SOLM-2018 database

Contact details

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

Colin Greenstreet
Community organiser,
Signs of Literacy

Email:
m.hailwood@bristol.ac.uk
colin.greenstreet@gmail.com

Weblinks:
http://signsofliteracy.org
http://marinelives.org
http://chronoscopic.org

GitHub:
https://github/Signsofliteracy/Signoff

Twitter:
Marinelivesorg

Working with: The National Archives, kaggle Competitions, iif, PICTURAE

Discussion

We need your input today at #dhc2018

	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)				+
Q3	Archives & AI Sheffield DH Congress	Kaggle competition Huntington library visit				**Machine learning enabled manuscript archivists** **Image processing** **Natural Language Programming** **Key word spotting** **Controlled vocabularies** **Programmable decision rules** **IIIF visibility**
Q4	Kaggle visit Stanford University visit Transkribus Vienna users conference					
	SOLM-2018	**SOLM-2019**	**SOLM-2020**	**SOLM-2021**	**SOLM-2022**	**SOLM-2023**

We need your input today at #dhc2018

Collection structure

- Should our SOLM-2018 data collection consist of full pages with digital bounding boxes on the signoffs? Or full pages and separate digital signoff snippets?
- Should we include all digital pages from which metadata is sourced and linked to the signoffs?

Collection character

- How should we think about "diversity" of images & metadata within our SOLM data collection?
  - e.g. mix of markes vs. initials vs. signatures; high pixel vs. low pixel; full image vs occluded vs. underlying curvature
- How can we increase the diversity of our manuscript sources?
  - Document type (ecclesiastical vs secular; legal vs. non-legal; parish vs. regional or national; English vs. Dutch vs. Swedish vs. French vs. Spanish)
  - Time period (C1t6th; C17th; C18th; other?)

Sample size

- How big should our sub-collections be for single occupations? [n=60; n=600; n= 6,000?...]
- Should minimum occupational sub-collection size be further constrained by year of signoff and/or place of residence and/or place of origin?
  - e.g. "**n must be => 60 for sailors living in Wapping born in the decade 1620-29?"**

Backup material

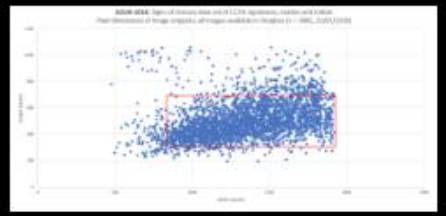
Project portfolio

MarineLives

Maphackathon

Signs of Literacy

SOLM-2018



EM Textiles, Garments & Dyestuffs Glossary

EM Maritime & Mercantile Gazetteer

http://www.chronoscopic.org

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

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

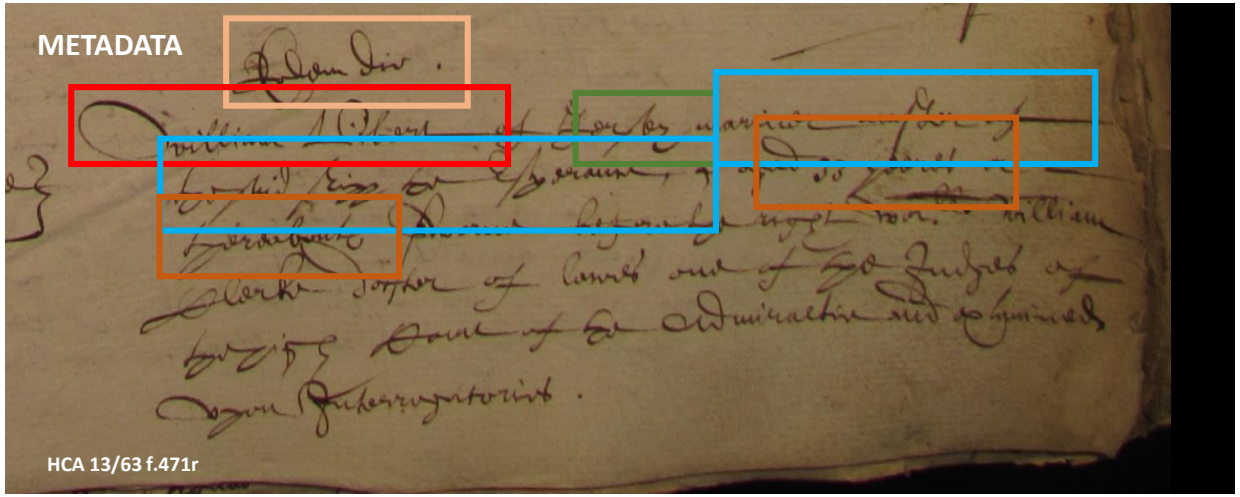
Source: David Cressy, Literacy and the Social Order (1980), (Cambridge, 1980), pp.132-133

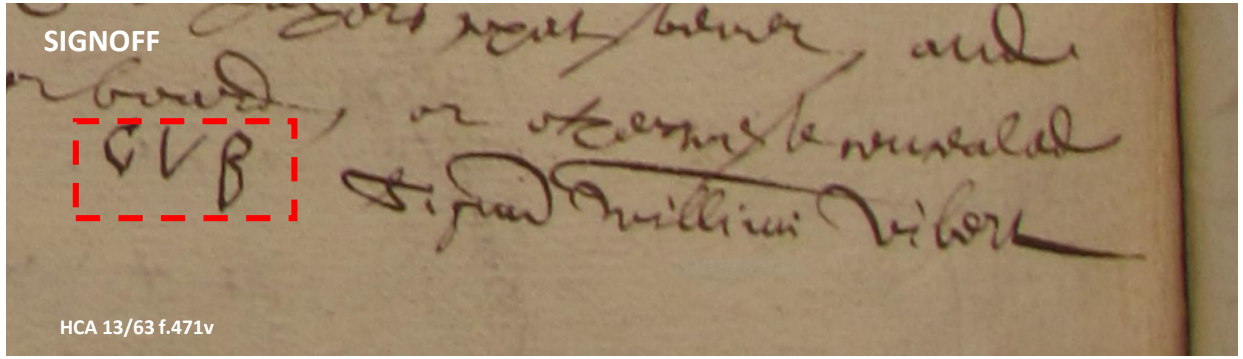
Table 7.3 *Illiteracy of tradesmen and craftsmen in the dioceses of Durham, Exeter, Norwich and London, 1560–1730*

	Durham/ Northumberland			Devon/Cornwall			Norfolk/Suffolk			Essex/Herts.			London/Middlesex		
Decade	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

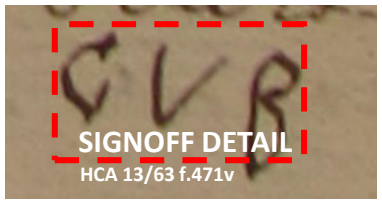
Source: David Cressy, Literacy and the Social Order (1980), (Cambridge, 1980), p.146

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 29th 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 marke. 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.

Source: http://www.marinelives.org/wiki/HCA_13/63_f.471r_Annotate; http://www.marinelives.org/wiki/HCA_13/63_f.471v_Annotate

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

Source: SOLM-2018 database